A SYSTEMATIC REVIEW AND META-ANALYSIS ASSESSING THE RELATIVE EFFICACY OF IMMUNE CHECKPOINT INHIBITORS BASED ON PD-L1 EXPRESSION LEVELS

Kathy Kwiatkowski

Submitted to the faculty of the University Graduate School
in partial fulfillment of the requirements
for the degree
Doctor of Philosophy
in the Richard M. Fairbanks School of Public Health,
Indiana University

October 2020

Accepted by the Graduate Faculty of Indiana University, in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

Doctoral Committee

Jiali Han, PhD, Co-Chair

Yiqing Song, MD, ScD, Co-Chair

September 2, 2020

Xin Li, ScD

Wanzu Tu, PhD

Acknowledgement

I would like to thank my husband Josh and my parents Phil and Sherry for never letting me give up during this journey. I want to thank my PhD committee for their advice and guidance during the dissertation process. This has been transformational, and I am grateful for the knowledge and skills gained through this process.

Kathy Kwiatkowski

A SYSTEMATIC REVIEW AND META-ANALYSIS ASSESSING THE RELATIVE EFFICACY OF IMMUNE CHECKPOINT INHIBITORS BASED ON PD-L1 EXPRESSION LEVELS

Purpose: The purpose was to comprehensively assess the impact of PD-L1 expression on the efficacy of immune checkpoint inhibitors on Overall Survival (OS) and Progression-Free Survival (PFS).

Methods: A systematic literature search and review was conducted through June 2019. I searched all eligible randomized controlled trials comparing PD-1/PD-L1 monotherapy to an active comparator in adult patients with advanced cancer across multiple tumor types. The Cochrane risk-of-bias tool was used to assess trial quality. A random-effects model was used for the meta-analysis. Heterogeneity was assessed using Cochran Q statistic and I^2 test. Publication bias was assessed by visual inspection of a funnel plot and Begg's test.

Results: I identified and included 23 trials involving 14,434 participants. When stratifying PD-L1 positive (+) and negative (-) patients using varying thresholds of expression, a significant group difference was observed at PD-L1 $\geq$1% (p=0.04; PD-L1(+): HR, 0.72; 95% CI, 0.65-0.79; PD-L1(-): HR,0.83; 95% CI, 0.75-0.91), at PD-L1 $\geq$10% (p=0.02; PD-L1(+): HR,0.50; 95% CI, 0.38-0.62; PD-L1 (-): HR, 0.74; 95% CI, 0.57-0.90) and at PD-L1$\geq$50% (p=0.01; PD-L1(+): HR,0.59; 95% CI, 0.51-0.68; PD-L1(-): HR, 0.93; 95% CI, 0.71-1.15). Across tumor types, both PD-L1(+) and PD-L1(-) patients treated with an immunotherapy had improved OS compared with patients receiving standard care therapies. A PFS benefit was observed and favored patients

treated with a PD-1/PD-L1 inhibitor versus standard of care. However, there was significant heterogeneity and the benefit on PFS was not statistically significant between PD-L1(+) and PD-L1(-) groups using varying cut-off levels of PD-L1 expression. No differences between sub-groups of interest including median follow-up time, type of inhibitor, and line of therapy for either PD-L1(+) or PD-L1(-) patients at 1% cut-off were identified.

Conclusion: This study supports the use of PD-L1 as a predictive biomarker of improved response to immunotherapies. As thresholds increase and specifically above the 10% PD-L1 expression threshold, patients who were positive for PD-L1 appeared to have better OS compared to those who were negative for PD-L1. Further investigation is needed to assess the clinical usefulness of PD-L1 at various expression levels with improved technologies that have the potential to enhance assay accuracy and precision.

Jiali Han, PhD, Co-Chair

Yiqing Song, MD, ScD, Co-Chair

Table of Contents

List of Tables

List of Figures

List of Abbreviations

Abbreviation	Definition
CI	confidence interval
DCF	data collection form
GRADE	Grading of Recommendations, Assessment, Development and Evaluations
HR	hazard ratio
IHC	immunohistochemistry
mAB	monoclonal antibody
NSCLC	non-small cell lung cancer
OS	overall survival
PD-1	programmed cell death protein 1
PD-L1	programmed death ligand 1
PD-L1(+)	PD-L1 expression positive
PD-L1(-)	PD-L1 expression negative
PFS	progression free survival
TMB	tumor mutational load or burden
TNBC	triple-negative breast cancer
TPS	tumor proportion score

Introduction

Cancer continues to be a significant cause of human disease, with an estimated 1.8 million new cases diagnosed in the United States and over 600,000 deaths in the U.S. in 2020 (Society, 2020). As cancer is the second leading cause of U.S. deaths, identifying improved treatment methods is critical to reducing the extreme burden placed on patients, their families, health care systems, and society (CDC, 2016). Within the last five years, cancer immunotherapy has emerged as a preferred treatment, demonstrating remarkable results in otherwise difficult-to-treat cancer types.

Immunotherapy is a type of cancer treatment that leverages the body's natural defenses to slow or stop the growth and spread of tumor cells. It also helps the body's immune system kill tumor cells more effectively (Cancer.net, 2019). Resistance to immunotherapy is believed to be due to tumor activation of immune-checkpoint pathways, which down-modulate immune function and prevent the rejection of cancer cells as foreign (Pardoll, 2012). Blockade of these immune checkpoints with antibody therapies has been successful in some tumors, enhancing both progression-free survival (PFS) and overall survival (OS) (Pardoll, 2012). There are multiple types of immunotherapies, and the focus of the current work is monoclonal antibodies, specifically programmed cell death protein 1 (PD-1) and programmed death ligand 1 (PD-L1) checkpoint inhibitors.

Five PD-1/PD-L1 inhibitors are currently approved by the US Food and Drug Administration for the treatment of cancer, including anti-PD-1 antibodies (pembrolizumab and nivolumab) and anti-PD-L1 antibodies (atezolizumab, avelumab and durvalumab). As outlined in Table 1, approvals span several major cancer types (e.g.,

melanoma, non-small cell lung cancer (NSCLC), renal, urothelial, and microsatellite instability-high colorectal cancer) and specific indications for use (Genentech, 2019; Merck & Co, 2020; Pharmaceuiticals, 2020; Serono, 2019; Squibb, 2020). However, despite intensive investigation, checkpoint inhibitors have shown limited activity and have not been approved in the treatment of many other cancer types including prostate, ovarian, pancreatic, and brain.

Table 1

FDA Approved Checkpoint Inhibitors

Drug	Target	Cancer Types
KEYTRUDA *pembrolizumab*	PD-1	Cervical, Colorectal, Endometrial, Esophageal, Gastric, Head and Neck, Hepatocellular Carcinoma, Hodgkin Lymphoma, Melanoma, Merkel Cell, MSI-H, NSCLC, Primary Mediastinal Large B-cell Lymphoma, Renal Cell Carcinoma, SCLC, TMB-H, Urothelial
OPDIVO *nivolumab*	PD-1	Colorectal carcinoma, head and Neck Squamous Cell, Hepatocellular Carcinoma, Hodgkin Lymphoma, Melanoma, NSCLC, Renal Cell Carcinoma, Urothelial Carcinoma
TECENTRIQ *atezolizumab*	PD-L1	NSCLC, SCLC, Triple-Negative Breast Cancer, Urothelial Carcinoma
BAVENCIO *avelumab*	PD-L1	Merkel Cell Carcinoma, Urothelial Carcinoma
IMFINZI *durvalumab*	PD-L1	NSCLC, Urothelial Carcinoma

Notably, while there has been tangible improvement in survival compared to chemotherapy, only a subset of patients within each approved indication respond to checkpoint therapy. Also, responses to checkpoint inhibitors have been reported outside of approved indications, suggesting the potential for new, biomarker-defined indications (Carbone, Reck, Paz-Ares, Creelan, Horn, Steins, Felip, van den Heuvel, Ciuleanu,

Badin, Ready, Hiltermann, Nair, Juergens, Peters, Minenza, Wrangle, Rodriguez-Abreu, Borghaei, Blumenschein, Villaruz, Havel, Krejci, Corral Jaime, Chang, Geese, Bhagavatheeswaran, Chen, Socinski, et al., 2017; Panda et al., 2018). Thus, identifying and deploying predictive response biomarkers is critical, both to make informed treatment decisions in approved indications and potentially to support indication expansion (Gibney, Weiner, & Atkins, 2016).

Several biomarkers have been explored including PD-L1 expression (by immunohistochemistry), tumor-infiltrating lymphocytes (such as effector CD8+ T-cells), T-cell receptor clonality, tumor mutational load or burden (TMB), microsatellite instability status, peripheral blood markers, immune gene signatures, and multiplex immunohistochemistry (Gibney et al., 2016). The most well-studied biomarker is PD-L1 expression, which is approved as a companion or complementary diagnostic for multiple checkpoint inhibitors. While PD-L1 expression enriches for response in some indications, it is not a perfect biomarker, with many biomarker-positive patients exhibiting little treatment response and biomarker-negative patients exhibiting substantial response (Borghaei, Paz-Ares, Horn, Spigel, Steins, Ready, Chow, Vokes, Felip, Holgado, Barlesi, et al., 2015; Brahmer, Reckamp, Baas, Crinò, et al., 2015; Garon et al., 2015; Larkin, Chiarion-Sileni, Gonzalez, Grob, Cowey, Lao, Schadendorf, Dummer, Smylie, Rutkowski, Ferrucci, et al., 2015; Mahoney & Atkins, 2014).

Likewise, multiple antibodies, staining protocols, and evaluation methodologies are utilized; e.g. some approaches consider PD-L1 expression only on tumor cells, while others consider both tumor and immune-cell expression. Similarly, the use of biomarkers beyond PD-L1 to identify patient subgroups likely to respond to checkpoint inhibitors or

who will have elevated risk of off-target effects (such as development of an autoimmune disease), has not yet identified a conclusive patient stratification biomarker (Gibney et al., 2016; Topalian, Taube, Anders, & Pardoll, 2016).

A systematic review and meta-analysis comprehensively assessing the relative efficacy of immunotherapies across PD-L1 expression levels across tumor types may help us better understand responder status and improve patient selection for these drugs. The hypothesis is that across tumor types, elevated PD-L1 expression levels predict an increased clinical benefit with exposure to PD-1/PD-L1 inhibitors.

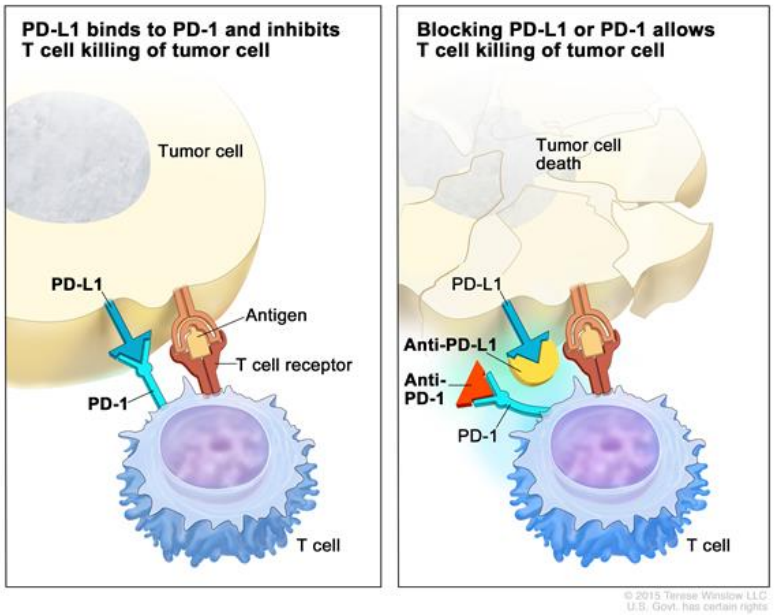
Anti-PD-1/PD-L1 Monoclonal Antibody Mechanism of Action

Scientists have been working for decades to identify immune mechanisms to stimulate a response to tumor cells. This has been mostly unsuccessful. The purpose of T cells is to destroy infected cells and signal to others to support the immune response. However, the systems or pathways that regulate the function of T lymphocytes (aka immune checkpoints) that exist to control excessive immune activation also seem to be a means for tumors to evade the immune system (Mahoney, Freeman, & McDermott, 2015).

PD-1 is a protein that indirectly effects cell death by interactions with its ligands (PD-L1 and PD-L1) (Figure 1). This signaling inhibits or inactivates T cells, which leads to reduced proliferation, cytokine production, and cell death (Chen, Irving, & Hodi, 2012; Mahoney et al., 2015). Tumor cells express PD-L1, which attaches to PD-1 and B7-1 receptors on T cells and thereby reduce the immune response (Chen et al., 2012; Keir, Butte, Freeman, & Sharpe, 2008).

Figure 1

PD-L1 and PD-1 Mechanism of Action



Note. Adapted with permission from the National Cancer Institute © 2015 Terese Winslow LLC, U.S. Govt.

Multiple tumors have elevated expression of PD-L1, including squamous cell carcinoma, colon adenocarcinoma, NSCLC, and breast adenocarcinoma; however, it is variable (Mahoney et al., 2015; Patel & Kurzrock, 2015). The recent approach with monoclonal antibodies (mAB) is to block the PD-L1 protein so tumor cells are unable to inactivate T cells (Chen et al., 2012). All the approved drugs in the anti-PD1/PD-L1 antibody category (Table 1) target the binding of PD-L1 to PD-1, but as Mahoney & Atkins note in their review of the affinity and isotype of the antibody, there are likely differences in clinical benefit (Mahoney & Atkins, 2014).

Pembrolizumab has the highest affinity for PD-1 of any PD-1 inhibitor, which may result in the antibody's ability to be effective at lower dosage or after prolonged

periods of stopping therapy (Mahoney et al., 2015). Further support for this notion was reported by De Sousa Linhares and colleagues in 2019 (De Sousa Linhares et al., 2019). They evaluated the efficacy of the five clinically used PD-1 inhibitors using a functional assay (T-cell reporter platform) and found that *in vitro* pembrolizumab is slightly more effective at blocking PD-1 than nivolumab, and that PD-1 antibodies are superior to PD-L1 antibodies in blocking PD-1 signaling (De Sousa Linhares et al., 2019).

Clinical Efficacy of PD-1 / PD-L1 Inhibitors

PD-L1/PD-1 inhibitors have revolutionized cancer care and their associated regulatory approvals have led to rapid adoption in clinical practice. Across five solid tumor types, Weng and colleagues observed in a recent meta-analysis that use of an immunotherapy reduced the risk of death by 31% (hazard ratio [HR]=0.69; 95% CI 0.64–0.74; P<0.00001) (Weng, Peng, Hu, Yao, & Song, 2018).

As of February 2020, there were 53 approved indications by FDA across the five PD-1/PD-L1 inhibitors (Appendix A) (Genentech, 2019; Merck & Co, 2020; Pharmaceuiticals, 2020; Serono, 2019; Squibb, 2020). All are approved in the advanced/refractory/metastatic setting, most indications are for combination therapy, and 16 are for monotherapy. Twenty-three of the approved labels are indicated for the first-line setting.

Notably, while improvement in survival compared to chemotherapy is tangible, only a subset of patients within each approved indication respond to checkpoint therapy. Thus, identification of predictive biomarkers has been a key focus to better understand who may benefit from immunotherapy.

PD-L1 as a Biomarker of Response

PD-L1 is perhaps the most-well studied biomarker and a reasonable choice given the mechanism of action of this type of inhibitor. Additionally, PD-L1 is well expressed across tumor types, though at variable levels. While cervical cancers and sarcomas are noted to have lower expression of PD-L1 (12-29%), others tumor types have a large percentage of samples expressing PD-L1 such as thymic (88-100%) or NSCLC with considerably variable estimates (20-95%) (Patel & Kurzrock, 2015). PD-L1 is routinely ordered by physicians to support immunotherapy decision-making, and PD-L1 positivity is currently required in a minority of approved drug labels.

Clinically, differential PD-L1 expression has been observed to enrich the response in some indications and is included as a companion diagnostic in a minority of drug labels. For example, Liu reported that in his meta-analysis across some solid tumors, OS improved in both PD-L1 positive (PD-L1(+)) patients (1% cut-off) (HR, 0.65; 95% CI, 0.60-0.70) and PD-L1 negative (PD-L1(-)) patients (HR, 0.82; 95% CI ,0.74-0.91) (Liu et al., 2019). This study had limitations including use of trial-level estimates, only patients with good performance status (which may have over-estimated treatment effects) included, moderate heterogeneity in the PD-L1 (-) group, and across only six tumor types (Liu et al., 2019).

PD-L1 is not a perfect biomarker, as many biomarker-positive patients exhibit little treatment response and negative patients frequently exhibit substantial response (Borghaei, Paz-Ares, Horn, Spigel, Steins, Ready, Chow, Vokes, Felip, Holgado, Barlesi, et al., 2015; Brahmer, Reckamp, Baas, CrinÃ², et al., 2015; Garon et al., 2015; Larkin, Chiarion-Sileni, Gonzalez, Grob, Cowey, Lao, Schadendorf, Dummer, Smylie,

Rutkowski, Ferrucci, et al., 2015; Mahoney & Atkins, 2014). Multiple factors may explain these discordant findings, including PD-L1 tests and their measurement variability, aspects of tumor-specific biology and heterogeneity of expression, as well as tissue age and quality (Kerr et al., 2015; Remon, Chaput, & Planchard, 2016).

PD-L1 expression is typically measured by immunohistochemistry, though technical variations include differences in general methods, inclusion of immune cells versus tumor cells, antibodies, and cut-off definitions, all which may contribute to the variable results of clinical studies.

Literature Gap

If PD-L1 were a perfectly predictive biomarker, we would expect the majority of patients who test positive to respond to immunotherapy. Previous meta-analyses of the literature indicated a dose-response relationship between PD-L1 expression levels and OS benefit beginning at the 1% cut-off across multiple tumor types (Liu et al., 2019; Weng et al., 2018).

In contrast, Davis and colleagues recently reported a meta-analysis of only trials used in regulatory filings, identifying significant limitations of using PD-L1 expression as a predictive biomarker and concluding that "PD-L1 was predictive in only 28.9% of cases, and was either not predictive (53.3%) or not tested (17.8%) in the remaining cases"(Davis & Patel, 2019). The authors also noted that this may be an over-estimate, because they only included studies that resulted in positive regulatory filings, increasing the likelihood that negative results were missed (Davis & Patel, 2019).

Based on the vast differences in approved indications and conflicting results reported previously, it is reasonable to undertake a systematic review and meta-analysis

of the updated literature on randomized clinical trials to comprehensively evaluate the utility of PD-L1 as a predictive biomarker to identify patients likely to respond to immune therapies targeting the PD-1/PD-L1 checkpoint. The results may help to support clinical decision making, tailor immunotherapeutics, and improve outcomes.

Methods

A comprehensive systematic review and meta-analysis of randomized controlled trials was conducted to assess the impact of PD-L1 expression on the efficacy of immune checkpoint inhibitors on OS and PFS. An electronic literature search was conducted using PubMed, Embase, ClinicalTrials.gov, and the Cochrane controlled trials search, separately, from their inception until June 2019. Only English language studies were considered. Additionally, as recent studies may be unpublished, electronic searches of the two relevant conferences (ASCO and ESMO) were reviewed. Bibliographies of published meta-analyses and other randomized controlled trials involving immunotherapies were also reviewed.

The main keywords and search terms included pembrolizumab, nivolumab, atezolizumab, avelumab, durvalumab, checkpoint inhibitors, PD-1, PD-L1, and randomized controlled trial. The scope of controlled clinical trials was limited to studies which assessed PD-1 monotherapy compared to an active therapy regimen in the advanced or metastatic solid tumor setting. The specific eligibility criteria are described below.

Eligibility Criteria

Studies were included based on the following criteria:

1. Randomized controlled trial
2. PD-L1 status of included patients (age >18) must be available
3. PD-L1 measurement and thresholds must be described
4. Treatment with PD-L1 or PD-1 inhibitors irrespective of dosage and duration
5. Active comparator

6. Any advanced or metastatic solid tumor
7. Primary endpoint included OS or PFS with hazard ratios available
8. Publication was available in English

Studies were excluded based on the following criteria:

1. PD-L1 expression levels were not described as a sub-group or stratification variable
2. PD-L1 expression results were not assessed with regards to OS or PFS
3. Studies assessing combination therapy including an immunotherapy

Data Collection & Quality

A data-collection form (DCF) was developed in Microsoft Access and piloted with two independent reviewers. Discrepancies were resolved and revisions were completed prior to starting the systematic review. After final selection of full-text articles for analysis, one reviewer independently extracted information from each study. Data was quality checked by an independent third party.

The data collection form included the following variables: NCT number, primary author, study acronym, tumor type, trial phase, randomized controlled trial designation, sample size, PD-L1 stratification, follow-up duration, PD-L1 assay description, PD-L1 method, primary outcome measurement, study drug and dosage, comparator drug and dosage, schedule, PD-L1 threshold, median age, sex, ECOG status, race, number of previous therapies, and outcome measurements. As available the following data for outcome measurements were documented by drug, dosage, and PD-L1 status; PFS, OS, hazard ratios, and confidence intervals and objective response rates.

Additionally, as recent studies may be unpublished, electronic searches of the two relevant conferences (ASCO and ESMO) were reviewed to minimize publication bias.

Analysis Plan

Analysis Objectives

The hypothesis was that across tumor types, as PD-L1 expression levels increase, greater clinical benefit will be demonstrated with exposure to PD-1/PD-L1 inhibitors as compared to standard-of-care chemotherapy regimens. The primary analysis was the difference in OS hazard ratios between PD-L1(+) versus PD-L1(-) patients. The secondary analysis was the difference in PFS hazard ratios between PD-L1(+) versus PD-L1(-) patients.

Population & Subgroups

Inclusion and exclusion criteria are listed in the Methods section. However, the review includes randomized controlled trials comparing PD-1/PD-L1 monotherapy to an active comparator (chemotherapy) in adult patients with advanced cancer across multiple tumor types. Subgroup meta-analyses stratified by PD-L1 (+) and PD-L1(-) groups were performed to specifically address potential interactions of PD-L1 levels with these inhibitors on efficacy measures using the Cochran Q-Statistic. Sub-groups (assessed at 1% threshold) included immunotherapy inhibitor type (PD-1 vs PD-L1), median duration of follow-up ($\geq$18months vs <18months), and line of therapy (1st vs later). Two sensitivity analyses were completed, including 1) evaluating NSCLC studies separately and 2) removal of a single study that used PFS as the primary outcome instead of OS. There was insufficient representation of melanoma results to conduct a sensitivity analysis.

Endpoints

Hazard Ratios and their respective confidence intervals for OS and PFS were used and reported separately by PD-L1 expression status. PFS is typically defined as the time from randomization in a clinical trial to disease progression or death from any cause. OS is defined as the time from randomization to death from any cause.

Handling of Missing Values and Other Data Conventions

Variables are summarized for all studies with available data. For key outcome variables, the proportion of missing data (i.e., if only PD-L1 positive groups are included in a study) will be described to understand the extent to which there may be under-reporting or bias.

Assessment of Risk of Bias

The Cochrane risk-of-bias tool (RoB2) was used to assess the quality of included randomized controlled trials (Sterne JAC, 2019). The tool assessed data across five domains, including randomization process, deviations from intended interventions, missing outcome data, measurement of outcomes, and selection of reported results. These domains are evaluated using signaling questions, weighted and summarized as low, some, or high concern of bias. Publication bias was assessed by visual inspection of a funnel plot (Egger, BMJ 1997) and using Begg's test for funnel plot asymmetry (Begg and Mazumdar, Biometrics 1994). Risk of small study effect was assessed for both OS and PFS outcomes for PD-L1 studies included in 1% cut-off analyses.

Descriptive Analysis

Summary tables were created to describe study characteristics, including study name, author, tumor type, phase, PD-L1 stratification, therapy line, sample size, median

duration of follow-up, PD-L1 staining type, percentage of males, percentage of ECOG 0 or 1 patients included, percent white, drug names and dosage, PD-L1 expression levels, and the primary outcome. Categorical variables (PD-L1 stratification, 1st line therapy studies, median follow-up, PD-L1 staining type and cut-offs) were summarized using the number and percentage of studies falling into each category.

Statistical Procedures

Analyses performed

The meta-analysis assessed both OS and PFS endpoint and was performed for all included studies combining all tumor types. A random-effects model based on the restricted maximum likelihood method was performed to combine the HRs from individual trial data using the approach defined by Hardy and Thompson (Hardy & Thompson, 1996). Primary meta-analysis methods using Stata *metaan* commands are also described in Palmer and Sterne's updated journal collection (Palmer, 2016). The *meta summarize, subgroup command* was used to report individual effect sizes and the overall effect size, their confidence intervals, heterogeneity statistics including a test of group differences, and with the *meta forestplot command* as graphical display. The test of group differences or interaction aims to determine if there is a difference in effect sizes between the sub-groups and is expressed using *Q*. The *reml* command was used to denote the restricted maximum likelihood method. All meta-analyses were sub-grouped according to their PD-L1 positivity level to understand any difference between positive versus negative biomarker status at various cut-offs (1%, 5%, 10%, and 50%). Additional sub-group meta-analyses for OS were completed for all tumor types at the 1% PD-L1 positivity threshold with regard to type of drug (PD-L1 vs PD-1), 1st line vs. later study,

and median duration of follow-up (≥18months vs <18months) separately for PD-L1(+) and PD-L1(-) patients.

Between trial heterogeneity was assessed using Cochran *Q* statistic and the degree using I^2 test (Higgins & Thompson, Statistics in Medicine, 2002). The Cochrane *Q* test ($P < 0.05$) indicated a high level of statistical heterogeneity and the percentage of I^2 of 25%, 50%, and 75% corresponding to low, moderate, and high degrees of heterogeneity, respectively.

Sensitivity analyses were conducted within the OS analysis at 1% and 5% thresholds by removing one study with PFS as its primary endpoint. Additionally, analyses were conducted separately for NSCLC to better understand any differences in that specific tumor type.

Statistical Programs

Stata version 16 (StataCorp LP, College Station, TX, USA) was used for all statistical analyses. A two-tailed $P < 0.05$ was considered statistically significant unless specified otherwise.

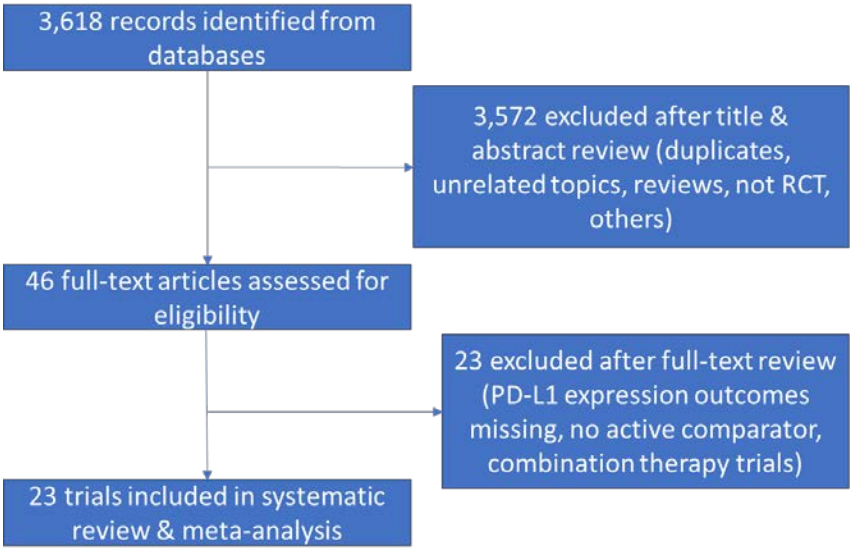
Results

Systematic Review Results

The search consisted of all studies published by June 2019 with complete results available. A total of 3,618 records were identified in the initial search across PubMed, Embase, and Cochrane Clinical Trials which included abstracts from ESMO and ASCO. Only 46 articles remained after preliminary screening of titles and abstracts, and 23 studies were finally included after full text review in the current systematic review and meta-analysis. Studies were removed for various reasons, including lack of overall survival endpoint data, study of combination therapy regimens, and absence of PD-L1-specific information. Across these 23 studies, 14,434 patients were represented. Ten studies focused on NSCLC, five on melanoma, and the remaining studies concerned head and neck, gastric, urothelial, and colorectal cancers. All studies were industry funded.

Figure 2

Article Selection Flowchart



The majority of studies were unblinded randomized controlled phase 3 trials, with the exception of two randomized phase 2 trials. Thirteen of the 24 studies were stratified by PD-L1 status. Seven studies were executed in the front-line setting, with the remainder being second-line or later. Studies averaged a total sample size of 682 patients with the majority gender being male (mean = 66%). For studies reporting ECOG status, the mean proportion of patients with a performance status of ECOG O was 40% and ECOG 1 mean was 58%. Only 10 studies reported race, and most of their patients were white (mean = 76%). Median age was 61.6 years and median follow-up time for endpoint assessment ranged from 7.3 months to 28 months.

Each randomized controlled trial compared a PD-1/PD-L1 monotherapy to standard-of-care options such as investigator choice of chemotherapy regimen or ipilimumab for patients with melanoma. Immunotherapies represented in this systematic review included two studies with avelumab (10mg/kg every 2 weeks [Q2W]), four studies with atezolizumab (1,200mg Q2W and Q3W), nine studies with nivolumab (3mg/kg Q2W), and eight studies with pembrolizumab (across multiple doses of 200mg, 2mg/kg, and 10mg/kg all Q3W schedule).

PD-L1 assessments were run on immunohistochemistry assays, mostly central laboratory assessments and included the following manufacturers: IHC 73-10 pharmDx, IHC 22C3 pharmDx, IHC Dako 28-8, and Ventana SP142. Of our 23 studies, 21 stained using tumor proportion score (TPS) versus combined proportion score (CPS).

Twenty-one studies included OS as a primary endpoint, and the other two used PFS. Six studies had co-primary endpoints of PFS and OS. The strategies for incorporation of PD-L1 cut-offs varied significantly between studies. Seventeen studies

included at least 1% as their PD-L1 cut-off and reported various endpoint analyses for the remaining endpoints. Over half of the studies stratified on PD-L1 variable within their studies; 16 of the studies allowed all-comers, while 7 studies excluded those that did not meet a specific PD-L1 cut-off. Individual study summary characteristics are listed in Table 2 and Table 3 (Bang et al., 2018; Barlesi et al., 2018; Bellmunt et al., 2017; Borghaei, Paz-Ares, Horn, Spigel, Steins, Ready, Chow, Vokes, Felip, Holgado, & et al., 2015; Brahmer, Reckamp, Baas, CrinÃ², et al., 2015; Carbone, Reck, Paz-Ares, Creelan, Horn, Steins, Felip, Van Den Heuvel, Ciuleanu, Badin, Ready, Hiltermann, Nair, Juergens, Peters, Minenza, Wrangle, Rodriguez-Abreu, Borghaei, Blumenschein, Villaruz, Havel, Krejci, Corral Jaime, Chang, Geese, Bhagavatheeswaran, Chen, & Socinski, 2017; Cohen et al., 2019; Eng et al., 2019; Fehrenbacher et al., 2016; Fehrenbacher et al., 2018; Ferris et al., 2018; Hamid et al., 2017; Herbst et al., 2019; Larkin, Chiarion-Sileni, Gonzalez, Grob, Cowey, Lao, Schadendorf, Dummer, Smylie, Rutkowski, & et al., 2015; Mok et al., 2019; Motzer et al., 2015; Powles et al., 2018; Reck et al., 2019; Reck et al., 2016; Robert et al., 2015; Schachter et al., 2017; Shitara et al., 2018; Weber et al., 2015; Wu et al., 2018).

Table 2

Overall Study Characteristics of the Included 23 Randomized Controlled Trials

ID	Study	Author	Tumor Type	Stratified by PD-L1?	Line	N	Follow-up (Median)	Primary Outcome	PD-L1 Cut-offs	PD-L1 (+) Only	Dosage	Schedule
1	JAVELIN Lung 200	Barlesi, F	NSCLC	Yes	2	792	18.9	OS	1%, 50%, 80%	Yes	10mg/kg	Q2W
											75mg/m2	Q3W
2	KEYNOTE-045	Bellmunt, J	Urothelial	No	2	542	14.1	OS, PFS	10%	Yes	200mg	Q3W
											multiple	multiple
3	CheckMate 078	Wu, Yi-Long	NSCLC	Yes	2	504	10.41	OS	1%	No	3mg/kg	Q2W
											75mg/m2	Q3W
4	JAVELIN Gastric 300	Bang, Y-J	Gastric	No	3	371	18	OS	1%	No	10mg/kg	Q2W
											Varied	Varied
5	CheckMate 141	Ferris, R.L.	Head & Neck	No	2	361	24.2	OS	1%, 5%, 10%	No	3mg/kg	Q2W
											Multiple	Q1W
6	CheckMate 017	Brahmer, J	NSCLC	No	2	272	11 (min)	OS	1%, 5%, 10%	No	3mg/kg	Q2W
											75mg/m2	Q3W
7	CheckMate 25	Motzer, R.J.	RCC	No	2	821	24	OS	1%, 5%	No	3mg/kg	Q2W
											10mg/kg	1 daily
8			NSCLC	Yes	1	541	13.5	PFS		Yes	3mg/kg	Q2W

	CheckMate 26	Carbone, D							1%, 5%, 50%		Multiple	Q3W
9	CheckMate 037	Weber, J	Melanoma	Yes	2	405	24	OS	5%	No	3mg/kg	Q2W
											Multiple	Q3W
10	CheckMate 57	Borghei, H.	NSCLC	No	2	582	17.2	OS	1%, 5%, 10%	No	3mg/kg	Q2W
											75mg/kg	Q3W
11	CheckMate 066	Robert, C	Melanoma	Yes	1	418	16.7	OS	5%	No	3mg/kg	Q2W
											1kmg/m2	Q3W
12	Checkmate 067	Larkin, J.	Melanoma	Yes	1	945	60 (min)	OS, PFS	1%, 5%, 10%	No	3mg/kg	Q2W
											3mg/kg	Q3W
13	Imblaze 370	Eng, C	Colorectal	No	2	363	7.3	OS	1%	No	1200mg	Q2W
											160mg	Daily
14	Imvigor211	Powles, T.	Urothelial	Yes	1	931	17.3	OS	5%	Yes	1200mg	Q3W
											multiple	Q3W
15	KEYNOTE-002	Hamid, O	Melanoma	No	2	540	28	OS	1%	No	2mg/kg	Q3W
											10mg/kg	Q3W
											Multiple	Multiple
16	KEYNOTE 006	Carlino, MS	Melanoma	Yes	1	834	22.9	OS, PFS	1%	No	10mg/kg	Q2W
											10mg/kg	Q3W
											3mg/kg	Q3W
17			NSCLC	Yes	3	1034	13.1	OS, PFS		Yes	multiple	Q3W

	KEYNOTE 010	Herbst, R							1%, 50%		75mg/kg	Q3W
18	KEYNOTE-024	Reck, M	NSCLC	No	1	305	25.2	PFS	50%	Yes	200mg	Q3W
											Multiple	Multiple
19	KEYNOTE-040	Cohen, E	Head & Neck	Yes	2	495	7.5	OS	50%	No	200mg	Q3W
											Multiple	Multiple
20	KEYNOTE-061	Shitara, K	gastro-esophogeal	Yes	2	592	8.5	OS, PFS	1%	No	200mg	Q3W
21	KEYNOTE-042	Mok, T	NSCLC	No	1	1274	12.8	OS	1%, 20%, 50%	Yes	200mg	Q3W
											multiple	multiple
22	OAK	Fehrenbacher, L	NSCLC	Yes	2	1225	28	OS	1%, 5%, 50%	No	1200mg	Q3W
											75mg/m2	Q3W
23	POPLAR	Fehrenbacher, L	NSCLC	Yes	2	287	14.8	OS	1%, 5%, 50%	No	1200mg	Q3W
											75mg/m2	Q3W

Table 3

Patient Demographics of the Included 23 Randomized Controlled Trials

ID	Study	Author	%Males	Males	Females	% ECOG 0	ECOG0	%ECOG1	ECOG1	%White	White	Age (Median)
1	JAVELIN Lung 200	Barlesi, F	68%	269	127	35%	144	65%	252	68%	273	64
				273	123		134		262		262	64
2	KEYNOTE-045	Bellmunt, J	74%	200	70	42%	119	56%	143	NR		67
				202	70		106		158			65
3	CheckMate 078	Wu, Yi-Long	79%	263	75	13%	47	86%	291	NR		60
				134	32		21		144			60
4	JAVELIN Gastric 300	Bang, Y-J	72%	140	45	35%	66	65%	119	64%	119	59
				127	59		62		124		117	61
5	CheckMate 141	Ferris, R.L.	83%	197	43	20%	49	78%	189	83%	196	59
				103	18		23		94		104	61
6	CheckMate 017	Brahmer, J	76%	111	24	24%	27	76%	106	93%	122	62
				97	40		37		100		130	64
7	CheckMate 25	Motzer, R.J.	75%	315	95	NR	0	NR		88%	353	62
				304	107		0				367	62
8	CheckMate 26	Carbone, D	61%	184	87	33%	85	66%	183	87%	228	63
				148	122		93		174		242	65
9	CheckMate 037	Weber, J	64%	176	96	61%	162	39%	110	NR		59
				85	48		84		48			62
10	CheckMate 57	Borghei, H.	55%	151	141	31%	84	69%	208	92%	267	61
				168	122		95		194		266	64
11	CheckMate 066	Robert, C	59%	121	89	64%	148	34%	60	NR		64
				125	83		121		84			66
12		Larkin, J.	64%	202	114	61%	238	39%	77	NR		59

	Checkmate 067			202	113		224		91			61
13	Imblaze 370	Eng, C	30%	59	31	24%	42	26%	48	40%	73	56
				51	39		45		45		71	59
14	Imvigor211	Powles, T.	77%	357	110	46%	218	54%	249	72%	335	67
				361	103		207		257		336	67
15	KEYNOTE-002	Hamid, O	61%	104	76	55%	98	45%	80	NR		62
				109	72		100		81			60
				114	64		98		81			63
16	KEYNOTE 006	Carlino, MS	60%	161	118	69%	196	31%	83	NR		31
				174	103		189		88			63
				162	116		188		90			62
17	KEYNOTE 010	Herbst, R	61%	425	265	34%	232	66%	454	NR		63
				209	134		116		224			62
18	KEYNOTE-024	Reck, M	61%	92	62	35%	54	65%	99	NR		64
				95	56		53		98			66
19	KEYNOTE-040	Cohen, E	83%	207	40	28%	71	72%	176	NR		60
				205	43		67		180			60
20	KEYNOTE-061	Shitara, K	69%	202	94	45%	127	55%	169	NR		62
				208	88		137		158			60
21	KEYNOTE-042	Mok, T	71%	450	187	31%	198	69%	439	NR		63
				452	158		192		445			64
22	OAK	Fehrenbacher, L	62%	379	234	37%	221	63%	392	71%	438	63
				379	233		234		378		432	64
23	POPLAR	Fehrenbacher, L	50%	93	76	49%	46	49%	45	NR		62
				51	67		96		97			62

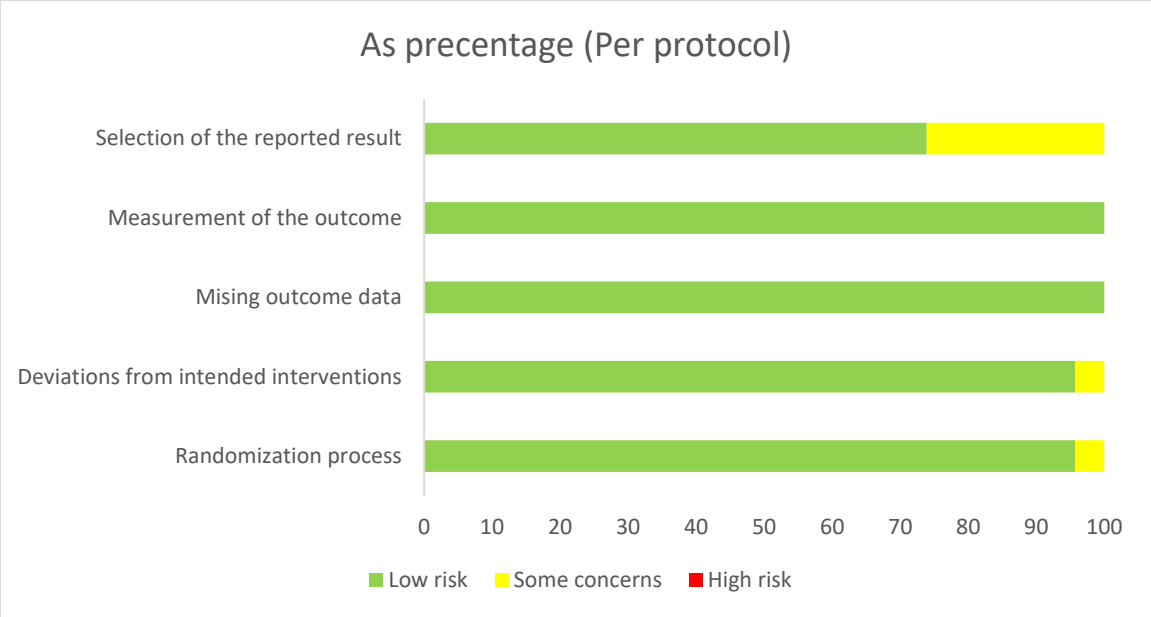
Regarding PD-L1 positivity, the reported PD-L1 expression levels varied across studies and their accompanying impact on clinical benefit. Among PD-L1(+) patients, 58% of studies at the 1% cut-off, 73% of studies at 5% cut-off, 60% of studies at 10% cut-off, and 78% of studies at the 50% cut-off showed statistically significant results, with OS benefit from a PD-1/PD-L1 monotherapy compared to standard of care. In contrast, regarding PD-L1(-) patients, 15% of studies at the 1% cut-off, 29% of studies at 5% cut-off, 25% of studies at 10% cut-off, and zero studies at the 50% cut-off (n=1) observed statistically significant results with a clinical benefit over the standard of care.

Risk of Bias Assessment

The Cochrane risk-of-bias 2 tool (RoB2) was used to assess the quality of the included randomized controlled trials. All the trials had low risk for bias due to inappropriate measures of outcomes and/or missing outcome data. The randomization process was reported in all trials, and there were no deviations from intended interventions in trials that may have affected outcomes. None of the trials showed a high probability of selection bias in the reported results. Overall, there was low risk of bias, given that 17 of the 23 studies were indicated as low risk based on the individual domains assessed. Figure 3 illustrates the percentage of protocols with risk assessments in each category.

Figure 3

Cochrane Risk of Bias (Per protocol) Summary by Domains



Though some tumor measurements were assessed by unblinded assessors, biased results were very unlikely. It was also observed that there were generally more dropouts prior to dosing in chemo groups versus immunotherapy groups, though this was also unlikely to affect study outcomes. Line output is available in Appendix B.

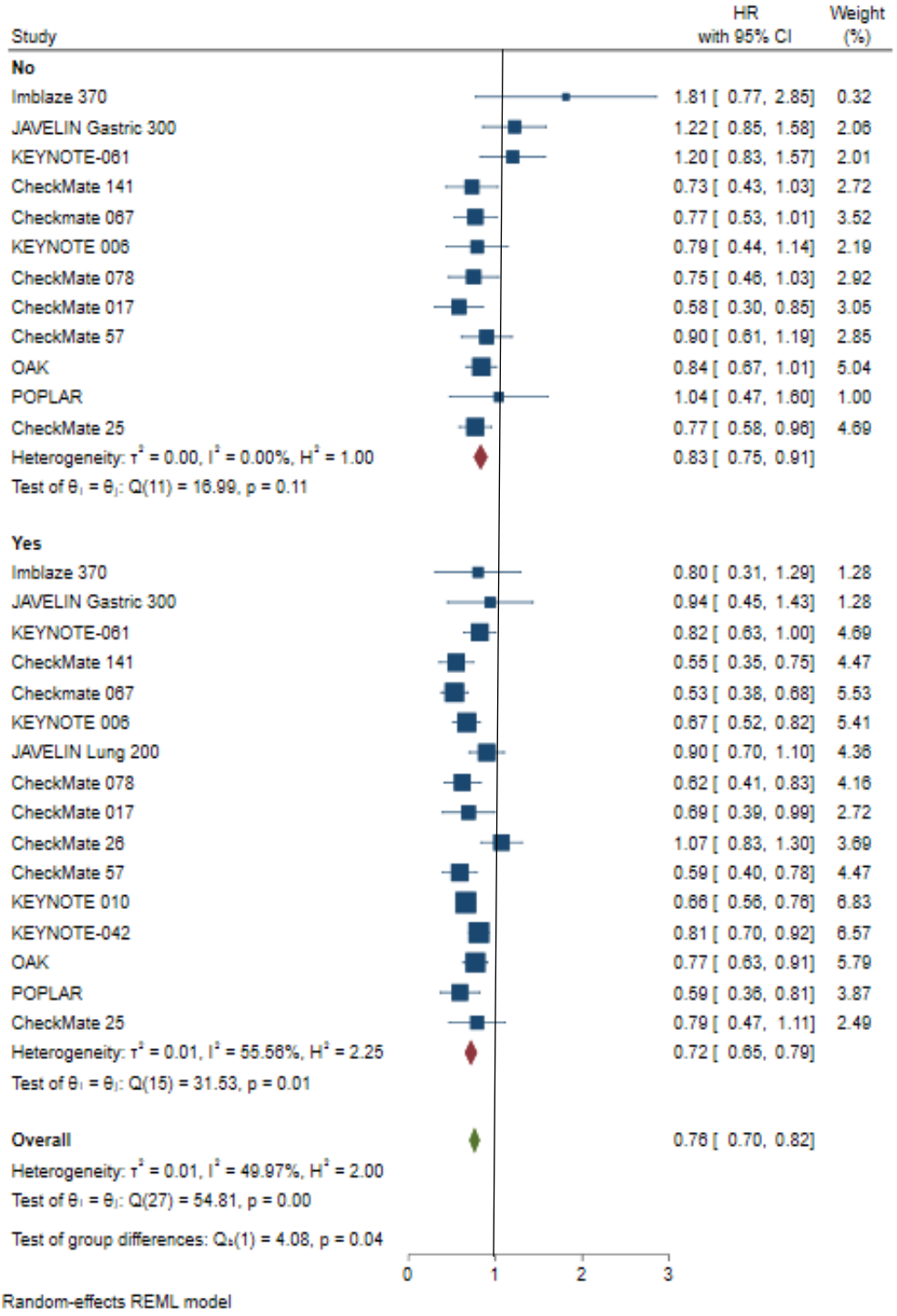
Meta-Analysis Results – Overall Survival (OS)

Using a random-effects model based on the restricted maximum likelihood method, HRs from individual trial data were combined. All meta-analyses were grouped according to their PD-L1 positivity level to detect any difference between positive versus negative biomarker status at various thresholds (1%, 5%, 10%, and 50%). Summary tables and forest plots are used to illustrate results from individual trials and the meta-analyses. All meta-analysis code, tables, and forest plot results are included in Appendix C.

Seventeen studies had available data for OS analysis at the 1% PD-L1 threshold across all solid tumors. The pooled estimate of PD-L1(+) effect at 1% threshold was 0.72 (95% CI 0.65-0.79). The pooled estimate of PD-L1(-) effect at 1% threshold was 0.83 (95% CI 0.75-0.91). A considerable level of heterogeneity was observed in the PD-L1(+) group (I^2 = 55.56%) and there was a modest yet statistically significant difference between the PD-L1(+) and PD-L1(-) groups (p=0.04). Figure 4 illustrates the Forest Plot of individual studies.

Figure 4

Overall Survival by PD-L1 Positivity using 1% Threshold Across All Tumors



Legend. No = Negative for PD-L1 at 1% threshold; Yes = Positive for PD-L1 at 1% threshold

No significant interaction was observed between PD-L1 positive and negative groups at the 5% PD-L1 threshold, although the survival benefit appears numerically better in PD-L1(+) patients than in PD-L1(-) patients. At the 10% PD-L1 threshold, the survival benefit is statistically significant and larger in PD-L1(+) patients than in PD-L1(-) patients (P =0.02). At the 50% threshold, any interpretation of group differences when stratifying between positive and negative groups should be approached with caution, as only one study presented data for negative patients; however, statistical significance was observed. Heterogeneity was not calculable in the negative group and was moderate in the positive group (I^2=32.29%). Table 4 summarizes the OS results using hazard ratios grouped by PD-L1 expression thresholds.

Table 4

Summary Table of Overall Survival Results by PD-L1 Expression Threshold

PD-L1 Threshold	Overall Survival HR (95% CI)			
	Studies (N)	PD-L1(+)	PD-L1(-)	Interaction
1%	17	0.72 (95% CI 0.65-0.79)	0.83 (95% CI 0.75-0.91)	p=0.04
5%	11	0.61 (95% CI 0.48-0.74)	0.76 (95% CI 0.57-0.95)	p=0.21
10%	5	0.50 (95% CI 0.38-0.62)	0.74 (95% CI 0.57-0.90)	p=0.02
50%	8	0.59 (95% CI 0.51-0.68)	0.93 (95% CI 0.71-1.15)	p=0.01

Subgroup Analysis – OS

Subgroup analyses were conducted at 1% PD-L1 expression thresholds for inhibitor type (PD-L1 versus PD-1), duration of follow-up (≥18mon vs <18mon), and line of therapy (1st vs later). No statistically significant differences were identified within

any of the sub-groups assessed (Table 5). Only one study used CPS vs TPS as the specific cell staining methodology, thus subgroup analyses were not performed.

Table 5

Sub-group Analysis for Overall Survival at 1% PD-L1 Threshold

Sub-Group	Overall Survival HR (95% CI)	
	PD-L1(+)	PD-L1(-)
PD-1 Inhibitor	0.70 (95% CI 0.61-0.79)	0.78 (95% CI 0.69-0.88)
PD-L1 Inhibitor	0.78 (95% CI 0.65-0.90)	1.05 (95% CI 0.76-1.35)
<18M Follow-up	0.78 (95% CI 0.65-0.90)	0.90 (95% CI 0.67-1.13)
≥18M Follow-up	0.70 (95% CI 0.61-0.79)	0.82 (95% CI 0.72-0.92)
1st Line	0.76 (95% CI 0.55-0.97)	0.78 (95% CI 0.58-0.98)
2nd or later	0.70 (95% CI 0.63-0.77)	0.86 (95% CI 0.74-0.98)

Sensitivity Analyses – OS

Sensitivity analyses were conducted at the PD-L1 thresholds of 1% and 5% with an attempt to explain heterogeneity in the positive group. CheckMate 026 was removed from the analysis, as it was the only study that had PFS as the primary endpoint, but also reported OS. This removal resulted in a reduction of 14% and 22% heterogeneity at the 1% and 5% thresholds, respectively *(Appendix C – Tables 11 and 12).*

Additionally, we limited the meta-analysis to NSCLC using the same random-effects analysis. There were insufficient data to assess OS within melanoma as a single tumor type. Table 13 in Appendix *C* illustrates an OS benefit seen in all NSCLC patients treated with immunotherapy vs standard of care; however, there was no statistically significant difference between PD-L1(+) and negative patients (p=0.47). A considerable level of heterogeneity was observed in the PD-L1(+) group (I^2 = 61.87%).

Meta-Analysis Results – Progression-Free Survival

No differences were observed between PD-L1 positive and negative groups at any of the thresholds analyzed using PFS as the outcome of interest, though considerable heterogeneity was identified at each threshold. A summary table of results is presented below.

Table 6

Summary Table of Progression-Free Survival Results by PD-L1 Expression Threshold

		Progression Free Survival HR (95% CI)		
PD-L1 Threshold	Studies (N)	PD-L1(+)	PD-L1(-)	Interaction
1%	17	0.81 (95% CI 0.67-0.95)	1.12 (95% CI 0.82-1.42)	p=0.06
5%	7	0.72 (95% CI 0.51-0.92)	0.84 (95% CI 0.40-1.28)	p=0.61
10%	10	0.57 (95% CI 0.42-0.73)	0.80 (95% CI 0.39-1.21)	p=0.32

Due to the significant heterogeneity across studies and lack of significance in differential response across positive and negative PD-L1 groups, PFS sub-group analyses were not attempted.

Publication Bias

Publication bias for Overall Survival (1% cut-off for determining PD-L1 (+) and PD-L1 (-) groups) was assessed by visual inspection of a funnel plot and Begg's test for symmetry. Statistically significant publication bias was observed (Prob > |z| = 0.05) for PD-L1 (-) results. However, no statistically significant publication bias was observed for PD-L1 (+) groups (Prob > |z| = 0.39). Funnel plots are below in Figures 5 and 6.

Figure 5

Publication Bias Funnel Plot using Overall Survival Outcome at 1% PD-L1 threshold for negative patients

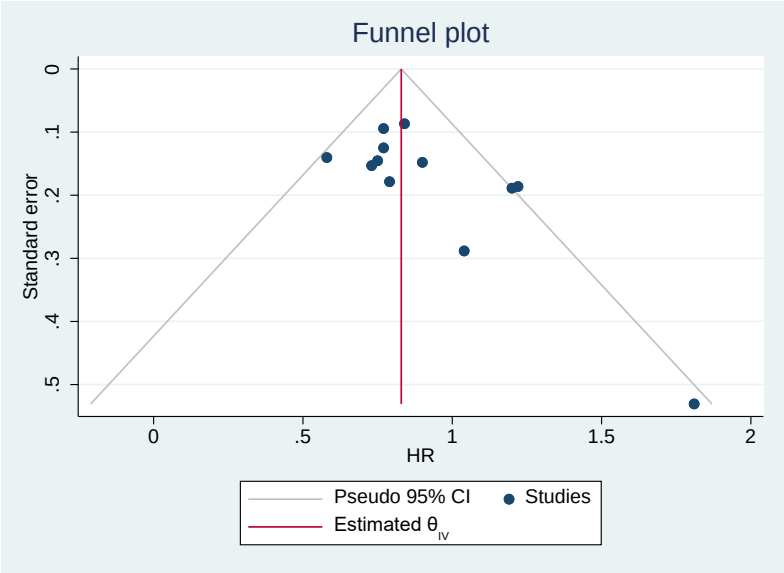
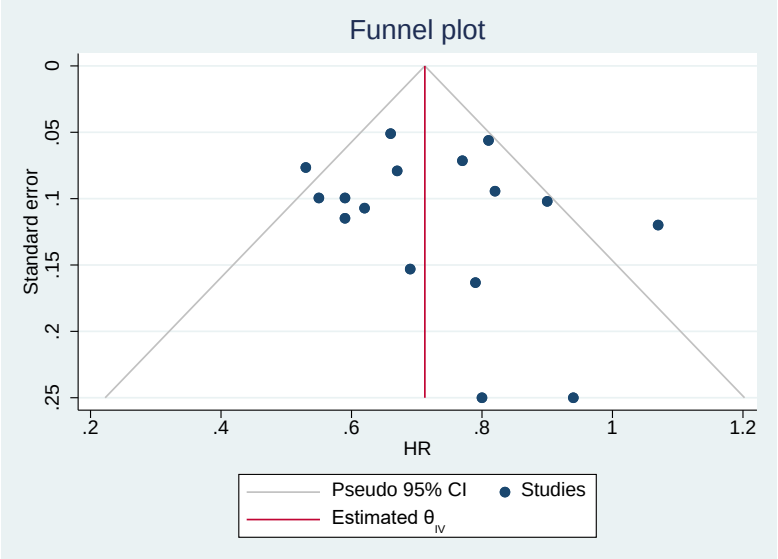


Figure 6

Publication Bias Funnel Plot using Overall Survival Outcome at 1% PD-L1 threshold for positive patients



Publication bias for PFS (1% cut-off for both PD-L1 positive and negative groups) was also assessed by visual inspection of a funnel plot and Begg's test. Statistically significant publication bias was found (Prob > |z| = 0.01) for PD-L1 (-) results. However, no statistically significant publication bias was observed for PD-L1 (+) groups (Prob > |z| = 0.19). Funnel plots are below in Figures 7 and 8.

Figure 7

Publication Bias Funnel Plot using PFS Outcome at 1% PD-L1 threshold for negative patients

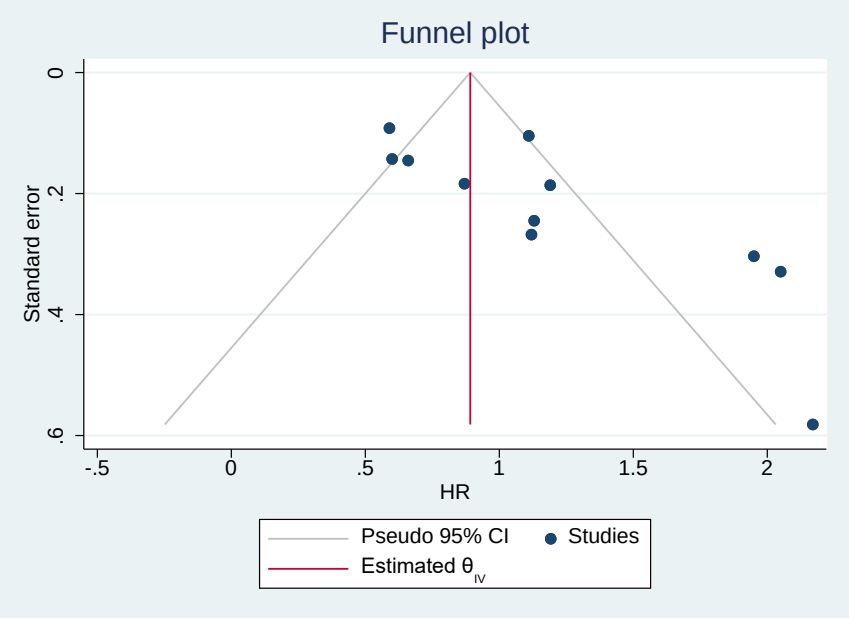
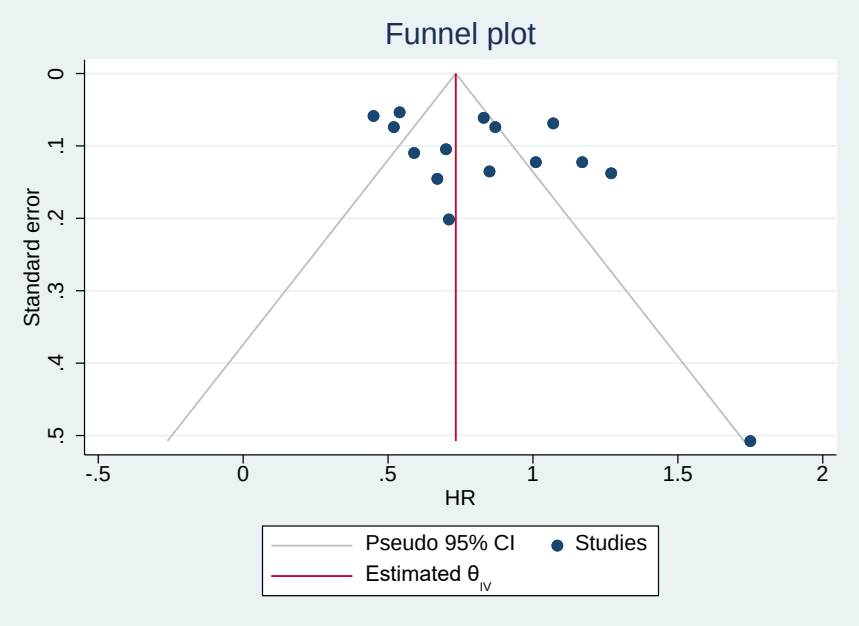


Figure 8

Publication Bias Funnel Plot using PFS Outcome at 1% PD-L1 threshold for positive patients



Discussion

The results of this study support the use of PD-L1 as a biomarker of improved response to immunotherapies. As thresholds increase and specifically above the 10% PD-L1 expression threshold, patients who are positive for PD-L1 appeared to have better OS compared to negative PD-L1 patients at the 10% threshold. This also supports the majority of the findings in the literature across tumor types, i.e., both PD-L1(+) and PD-L1(-) patients treated with PD-1/PD-L1 monotherapy have an improved OS over standard-of-care therapies (Genentech, 2019; Liu et al., 2019; Merck & Co, 2020; Pharmaceuiticals, 2020; Serono, 2019; Squibb, 2020; Weng et al., 2018; Zhang et al., 2016). The clinical benefit was seen in PFS as well for immunotherapies versus standard of care, though no interaction effect for PD-L1 levels was observed.

Using the 5% PD-L1 threshold to separate groups, no statistically significant difference was observed, though the groups were different numerically. While a significant difference was observed between PD-L1 positive and negative groups at 1% for all tumors, a moderate level of heterogeneity among studies was identified in the positive group; thus, all differences cannot be attributed to the PD-L1 group. Furthermore, no differences were identified between sub-groups of interest, including median follow-up time, type of inhibitor, and line of therapy for either PD-L1(+) or PD-L1(-) patients at 1% cut-off.

A sensitivity analysis was completed to assess potential sources of heterogeneity by removing the single study (CheckMate 26) that was powered for PFS outcome; however, that study also reported OS. This explained 22% of the heterogeneity within the PD-L1(+) group at the 5% level, and results trended towards positive group differences.

Within the 1% group, this sensitivity analysis explained 14% of the differences and further increased the statistical difference between the two groups.

PD-L1 expression is variable across tumor types however, an improved response to immunotherapy was seen across tumor types. Though there were insufficient studies to limit to melanoma, a sensitivity analysis limited to NSCLC revealed no difference between PD-L1 positive and negative groups, supporting the hypothesis of clinical benefit outside of this tumor type.

A PFS benefit was observed in PD-L1 positive patients at the thresholds analyzed and favored patients treated with a PD-1/PD-L1 inhibitor versus standard of care. However, there was significant heterogeneity, and no interaction between positive and negative groups was observed. Any interpretation of PFS analyses should be limited.

Immunohistochemistry (IHC) is the standard method for PD-L1 assessment. However, one recent study indicated a 100-fold difference in analytic sensitivity across reagents used for IHC, and another indicated an inter-rater reliability of assessing PD-L1 at the 1% level between 77% and 80% (Marchetti et al., 2017; Sompuram, Vani, Schaedle, Balasubramanian, & Bogen, 2018). This is clinically important, as low-level results may be affected by test manufacturer and individual pathology review. This also may contribute to the variability in findings at the 1% and to a lesser extent the 5% cut-off levels. Scoring methodologies in previous trials may also contribute to variability, as TPS included only tumor cells, and CPS included both tumor and immune cells. However, all but two of the studies included in this analysis used TPS as its staining and scoring methodology, so a formal analysis between TPS and CPS was not completed.

Measurement of PD-L1 expression could be improved by the use of next-generation sequencing technologies such as RNA-seq, to more reliably measure PD-L1 levels in tumor specimens. Conroy and colleagues recently used RNA-seq to assess analytical and clinical performance of detecting PD-L1 as compared to IHC and concluded it was comparable in performance and "had advantages of being amenable to standardization and avoidance of interpretation bias" (Conroy et al., 2019).

Though previous meta-analyses indicated a positive dose-response effect by PD-L1 expression thresholds, the study by Liu et al. assessed only nivolumab and involved 1,738 patients as compared to the current study, which includes over 14,000 (Liu et al., 2019). However, the study also had less variability in their testing methodologies, which may have limited bias and improved their ability to detect differences across all thresholds analyzed (Liu et al., 2019). Zhang and colleagues observed that PD-L1 positive patients had greater clinical benefit compared to PD-L1 negative patients when combining multiple thresholds in their sub-group analysis. The study included less than half the number of patients in the current study, as well as hematologic malignancies and patients treated with combination therapy (Zhang et al., 2016).

PD-L1 is now a routinely tested clinical biomarker. However, FDA approvals for PD-L1 specific indications are limited to 10 of the 53 approved, given the overall greater benefit of immunotherapy versus standard of care in advanced solid tumors. However, not all tumor types are approved for treatment with immunotherapy. PD-L1 is required only in indications held by pembrolizumab or atezolizumab within cervical, esophageal, gastric, head and neck, NSCLC, TNBC, and urothelial cancer. Given the improved response seen across tumor types at high levels of PD-L1 within the current study and

others, additional clinical trials are warranted in patients who have high PD-L1 expression, are outside of the approved indications, or who have very few therapeutic options for treatment of their advanced cancer.

The strengths of this study include its large sample size (over 14,000 patients) and up-to-date inclusion of all randomized clinical trials, as this is a fast-paced clinical development space. Another strength is the strict inclusion criteria of including only PD-1/PD-L1 monotherapies against an active comparator in the phase III setting, with 96% of studies powered for long-term OS as its primary endpoint. Combined with rigorous statistical methods, this study may provide relevant information to support further study of PD-L1 as a predictive biomarker in the clinical immunotherapy setting of patients with advanced solid tumors.

This systematic review also has limitations. Only one reviewer selected the full-text articles for inclusion, which may have biased selection. However, the articles selected were also reviewed against other meta-analyses, significant overlap was noted, and all data points for analysis were quality checked by a third party. Additionally, due to the considerable heterogeneity identified across all included trials or those in sub-groups, a random effects model was used. No significant publication bias was identified, although this specific analysis was underpowered, it suggests the usefulness of these results in evaluating the clinical impact of using PD-L1 as a predictive biomarker. Additionally, no trials included in this study reported co-factors that are likely contributors to PD-L1 variability, including pathologist variability, tissue age and quality, and specimen collection site. Further study is warranted to explore the impact of these potential co-factors on PD-L1 as a predictive biomarker of response to immunotherapy.

Using the GRADE (Grading of Recommendations, Assessment, Development and Evaluations) guidelines (Guyatt GH et al. BMJ, 2008) to rate the quality of the evidence for PD-L1 as a biomarker, which at higher thresholds indicate an improved response to immunotherapies, the strength of evidence is assessed to be high. Studies conducted to date resulting in the approvals of immunotherapies have been randomized controlled trials using active comparators with low risk of bias. While there is some heterogeneity across studies assessing different tumor types, measurement tools, and other pre-specified factors, the outcomes of interest and patient populations are similar with sufficient power to detect differences. Additionally, publication bias is estimated to be low, and there is evidence of a dose-response gradient as additional thresholds of PD-L1 expression are reached.

Conclusion

In conclusion, this systematic review and meta-analysis of randomized controlled trials support the use of PD-L1 as a predictive biomarker of improved response to immunotherapies. As thresholds increase, and specifically above the 10% PD-L1 expression threshold, patients who are positive for PD-L1 appeared to have better OS as compared to those negative for PD-L1. Further investigation is needed to assess the clinical usefulness of PD-L1 at various expression levels with different technologies that may have reduced random or systematic measurement error variability. Additional randomized trials, especially among patients with unapproved tumor types and higher levels ($\geq$10%) of PD-L1 expression, are warranted to provide further evidence regarding PD-L1 expression as a predictive biomarker for the use of PD-1/PD-L1 immune checkpoint inhibitors.

Appendix A

Brand Name	INN Name	Cancer Type	Indication Statement	PD-L1 Required?	In Combination With	Patient Population	Disease Stage	Line of Therapy	NCCN Category of Evidence	NCCN Cateory of Preferenc e
KEYTRUDA	pembrolizumab	Cervical Cancer	For the treatment of patients with recurrent or metastatic cervical cancer with disease progression on or after chemotherapy whose tumors express PD-L1 [Combined Positive Score (CPS) ≥1] as determined by an FDA-approved test	Yes	on or after chemotherapy	none	recurrent or metastatic	Second-line	Category 2A	Preferred
KEYTRUDA	pembrolizumab	Colorectal Cancer	For the treatment of adult and pediatric patients with unresectable or metastatic, microsatellite instability-high (MSI-H) or mismatch repair deficient colorectal cancer that has progressed following treatment with a fluoropyrimidine, oxaliplatin, and irinotecan	No	fluoropyrimidine, oxaliplatin, and irinotecan	Adult, pediatric	unresectable or metastatic	First-line	Category 2A	Preferred

OPDIVO	nivolumab	Colorectal Cancer	Adult and pediatric (12 years and older) patients with microsatellite instability-high (MSI-H) or mismatch repair deficient (dMMR) metastatic colorectal cancer that has progressed following treatment with a fluoropyrimidine, oxaliplatin, and irinotecan, as a single agent or in combination with ipilimumab	No	fluoropyrimidine, oxaliplatin, irinotecan, ipilimumab	12 years and older	microsatellite instability-high (MSI-H) or mismatch repair deficient (dMMR)	First-line	Category 2B	Alternative Preferred
KEYTRUDA	pembrolizumab	Endometrial Carcinoma	In combination with lenvatinib, for the treatment of patients with advanced endometrial carcinoma that is not MSI-H or dMMR, who have disease progression following prior systemic therapy and are not candidates for curative surgery or radiation	No	lenvatinib, prior systemic therapy and are not candidates for curative surgery or radiation	none	advanced endometrial carcinoma that is not MSI-H or dMMR	Not reported	Category 2A	Alternative Preferred
KEYTRUDA	pembrolizumab	Esophageal Cancer	For the treatment of patients with recurrent locally advanced or metastatic squamous cell carcinoma of the esophagus whose tumors express PD-L1 [Combined Positive Score (CPS) ≥10] as determined by an FDA-approved test, with disease progression after one or more prior lines of systemic therapy	Yes	after one or more prior lines of systemic therapy	none	recurrent locally advanced or metastatic	Second-line for CPS >10 or MSI-H or dMMR tumors. Third-line CPS>1	Category 2A	Preferred

KEYTRUDA	pembrolizumab	Gastric Cancer	For the treatment of patients with recurrent locally advanced or metastatic gastric or gastroesophageal junction adenocarcinoma whose tumors express PD-L1 [Combined Positive Score (CPS) ≥1] as determined by an FDA approved test, with disease progression on or after 2 or more prior lines of therapy including fluoropyrimidine- and platinum-containing chemotherapy and if appropriate, HER2/neu-targeted therapy	Yes	fluoropyrimidine- and platinum-containing chemotherapy and if appropriate, HER2/neu-targeted therapy	none	recurrent locally advanced or metastatic	Second-line for MSI-H or dMMR tumors. Third-line CPS>1	Category 2A	Preferred
KEYTRUDA	pembrolizumab	Head and Neck Squamous Cell Cancer (HNSCC)	In combination with platinum and FU for the first-line treatment of patients with metastatic or with unresectable, recurrent HNSCC	No	platinum and FU	none	metastatic or with unresectable, recurrent HNSCC	First-line	Category 2A	Preferred
KEYTRUDA	pembrolizumab	Head and Neck Squamous Cell Cancer (HNSCC)	As a single agent for the first-line treatment of patients with metastatic or with unresectable, recurrent HNSCC whose tumors express PD-L1 [Combined Positive Score (CPS) ≥1] as determined by an FDA-approved test	Yes	none	none	metastatic or with unresectable, recurrent HNSCC	First-line	Category 2A	Preferred

KEYTRUDA	pembrolizumab	Head and Neck Squamous Cell Cancer (HNSCC)	As a single agent for the treatment of patients with recurrent or metastatic HNSCC with disease progression on or after platinum-containing chemotherapy	No	on or after platinum-containing chemotherapy	none	recurrent or metastatic HNSCC	Second-line	Category 2A	Preferred
OPDIVO	nivolumab	Head and Neck Squamous Cell Cancer (HNSCC)	Patients with recurrent or metastatic squamous cell carcinoma of the head and neck with disease progression on or after a platinum-based therapy.	No	platinum-based therapy	none	recurrent or metastatic	Second-line	Category 2A	Preferred
KEYTRUDA	pembrolizumab	Hepatocellular Carcinoma (HCC)	For the treatment of patients with HCC who have been previously treated with sorafenib	No	sorafenib	none	none	Second-line	Category 2B	Alternative Preferred
OPDIVO	nivolumab	Hepatocellular Carcinoma (HCC)	Patients with hepatocellular carcinoma who have been previously treated with sorafenib	No	sorafenib	none	none	Second-line	Category 2A	Alternative Preferred
KEYTRUDA	pembrolizumab	Hodgkin Lymphoma	For the treatment of adult and pediatric patients with refractory cHL, or who have relapsed after 3 or more prior lines of therapy	No	who have relapsed after 3 or more prior lines of therapy	Adult, pediatric	refractory cHL	Third-line	Category 2A	Alternative Preferred
OPDIVO	nivolumab	Hodgkin Lymphoma	Adult patients with classical Hodgkin lymphoma that has relapsed or progressed after autologous hematopoietic stem cell transplantation (HSCT) and brentuximab vedotin	No	autologous hematopoietic stem cell transplantation (HSCT) and	Adult	relapsed or progressed	Second-line	Category 2A	Alternative Preferred

					brentuximab vedotin					
OPDIVO	nivolumab	Hodgkin Lymphoma	Adult patients with classical Hodgkin lymphoma that has relapsed or progressed after 3 or more lines of systemic therapy that includes autologous HSCT.	No	3 or more lines of systemic therapy that includes autologous HSCT.	Adult	relapsed or progressed	Third-line	Category 2A	Alternative Preferred
KEYTRUDA	pembrolizumab	Melanoma	For the adjuvant treatment of patients with melanoma with involvement of lymph node(s) following complete resection	No	none	none	involvement of lymph node(s) following complete resection	Second-line	Category 1	Preferred
KEYTRUDA	pembrolizumab	Melanoma	For the treatment of patients with unresectable or metastatic melanoma	No	none	none	unresectable or metastatic	First-line	Category 1	Preferred
OPDIVO	nivolumab	Melanoma	Patients with melanoma with lymph node involvement or metastatic disease who have undergone complete resection, in the adjuvant setting	No	none	none	lymph node invovlement or metastatic	Second-line	Category 1	Preferred
OPDIVO	nivolumab	Melanoma	Patients with unresectable or metastatic melanoma, as a single agent or in combination with ipilimumab	No	ipilimumab	none	unresectable or metastatic	First-line	Category 1	Preferred

BAVENCIO	avelumab	Merkel Cell Carcinoma (MCC)	Adults and pediatric patients 12 years and older with metastatic Merkel cell carcinoma (MCC)	No	none	Adults and pediatric patients 12 years and older	metastatic	Not reported	Category 2A	Preferred
KEYTRUDA	pembrolizumab	Merkel Cell Carcinoma (MCC)	For the treatment of adult and pediatric patients with recurrent locally advanced or metastatic Merkel cell carcinoma	No	none	Adult, pediatric	recurrent locally advanced or metastatic	Not reported	Category 2A	Preferred
KEYTRUDA	pembrolizumab	Microsatellite Instability-High Cancer	For the treatment of adult and pediatric patients with unresectable or metastatic, microsatellite instability-high (MSI-H) or mismatch repair deficient solid tumors that have progressed following prior treatment and who have no satisfactory alternative treatment options	No	following prior treatment and who have no satisfactory alternative treatment options	Adult, pediatric	unresectable or metastatic	Unknown	Category 2A	Unknown
IMFINZI	durvalumab	Non-Small Cell Lung Cancer (NSCLC)	Unresectable, Stage III non-small cell lung cancer (NSCLC) whose disease has not progressed following concurrent platinum-based chemotherapy and radiation therapy	No	none	none	unresectable, stage III	Second-line	Category 1	Preferred
KEYTRUDA	pembrolizumab	Non-Small Cell Lung	In combination with carboplatin and either paclitaxel or paclitaxel	No	carboplatin and either paclitaxel or	none	metastatic	First-line	Category 2A	Preferred

		Cancer (NSCLC)	protein-bound, as first-line treatment of patients with metastatic squamous NSCLC		paclitaxel protein-bound					
KEYTRUDA	pembrolizumab	Non-Small Cell Lung Cancer (NSCLC)	As a single agent for the first-line treatment of patients with NSCLC expressing PD-L1 [Tumor Proportion Score (TPS) ≥1%] as determined by an FDA-approved test, with no EGFR or ALK genomic tumor aberrations, and is metastatic.	Yes	none	none	metastatic	First-line	Category 1	Preferred
KEYTRUDA	pembrolizumab	Non-Small Cell Lung Cancer (NSCLC)	As a single agent for the treatment of patients with metastatic NSCLC whose tumors express PD-L1 (TPS ≥1%) as determined by an FDA-approved test, with disease progression on or after platinum-containing chemotherapy.	Yes	on or after platinum-containing chemotherapy.	none	metastatic	First-line	Category 1	Preferred
KEYTRUDA	pembrolizumab	Non-Small Cell Lung Cancer (NSCLC)	In combination with pemetrexed and platinum chemotherapy, as first-line treatment of patients with metastatic nonsquamous NSCLC, with no EGFR or ALK genomic tumor aberrations.	No	pemetrexed and platinum chemotherapy	none	metastatic nonsquamous NSCLC, with no EGFR or ALK genomic tumor aberrations	First-line	Category 1	Preferred

KEYTRUDA	pembrolizumab	Non-Small Cell Lung Cancer (NSCLC)	As a single agent for the first-line treatment of patients with NSCLC expressing PD-L1 [Tumor Proportion Score (TPS) ≥1%] as determined by an FDA-approved test, with no EGFR or ALK genomic tumor aberrations, and is stage III where patients are not candidates for surgical resection or definitive chemoradiation	Yes	none	none	stage III where patients are not candidates for surgical resection or definitive chemoradiation	First-line	Category 2A	Preferred
OPDIVO	nivolumab	Non-Small Cell Lung Cancer (NSCLC)	Patients with metastatic non-small cell lung cancer and progression on or after platinum-based chemotherapy	No	platinum-based chemotherapy	none	metastatic	First-line	Category 1	Preferred
OPDIVO	nivolumab	Non-Small Cell Lung Cancer (NSCLC)	Patients with metastatic small cell lung cancer with progression after platinum-based chemotherapy and at least one other line of therapy	No	platinum-based chemotherapy	none	metastatic	First-line	Category 1	Preferred
TECENTRIQ	atezolizumab	Non-Small Cell Lung Cancer (NSCLC)	for the treatment of adult patients with metastatic NSCLC who have disease progression during or following platinumcontaining chemotherapy	No	platinumcontaining chemotherapy	Adult	metastatic	First-line	Category 2A	Preferred
TECENTRIQ	atezolizumab	Non-Small Cell Lung	In combination with bevacizumab, paclitaxel, and carboplatin, for the first-line treatment of adult patients	No	bevacizumab, paclitaxel, and carboplatin	Adult	metastatic non-squamous NSCLC with no	First-line	Category 1	Preferred

		Cancer (NSCLC)	with metastatic non-squamous NSCLC with no EGFR or ALK genomic tumor aberrations				EGFR or ALK genomic tumor aberrations			
TECENTRIQ	atezolizumab	Non-Small Cell Lung Cancer (NSCLC)	in combination with paclitaxel protein-bound and carboplatin for the first-line treatment of adult patients with metastatic non-squamous NSCLC with no EGFR or ALK genomic tumor aberrations	No	paclitaxel protein-bound and carboplatin	Adult	metastatic non-squamous NSCLC with no EGFR or ALK genomic tumor aberrations	First-line	Category 1	Preferred
KEYTRUDA	pembrolizumab	Primary Mediastinal Large B-Cell Lymphoma (PMBCL)	For the treatment of adult and pediatric patients with refractory PMBCL, or who have relapsed after 2 or more prior lines of therapy	No	who have relapsed after 2 or more prior lines of therapy	Adult, pediatric	refractory PMBCL	Refractory	Category 2A	Not Reported
BAVENCIO	avelumab	Renal Cell Carcinoma (RCC)	First-line treatment, in combination with axitinib of patients with advanced renal cell carcinoma (RCC)	No	axitinib	none	advanced	First-line	Category 2A	Alternative Preferred
KEYTRUDA	pembrolizumab	Renal Cell Carcinoma (RCC)	In combination with axitinib, for the first-line treatment of patients with advanced RCC	No	axitinib	none	advanced	First-line	Category 2A	Preferred
OPDIVO	nivolumab	Renal Cell Carcinoma (RCC)	Patients with advanced renal cell carcinoma who have received prior anti-angiogenic therapy	No	prior anti-angiogenic therapy	none	advanced	Second-line	Category 2A	Preferred

OPDIVO	nivolumab	Renal Cell Carcinoma (RCC)	Patients with intermediate or poor risk, previously untreated advanced renal cell carcinoma, in combination with ipilimumab	No	ipilimumab	none	intermediate or poor risk	First-line	Category 2A	Alternative Preferred
KEYTRUDA	pembrolizumab	Small Cell Lung Cancer (SCLC)	For the treatment of patients with metastatic SCLC with disease progression on or after platinum-based chemotherapy and at least one other prior line of therapy	No	on or after platinum-conta on or after platinum-based chemotherapy and at least one other prior line of therapyining chemotherapy.	none	metastatic	Subsequent	Category 2A	Alternative Preferred
TECENTRIQ	atezolizumab	Small Cell Lung Cancer (SCLC)	In combination with carboplatin and etoposide, for the first-line treatment of adult patients with extensive-stage small cell lung cancer (ES-SCLC)	No	carboplatin and etoposide	Adult	extensive-stage small cell lung cancer (ES-SCLC)	First-line	Category 1	Preferred
TECENTRIQ	atezolizumab	Triple-Negative Breast Cancer (TNBC)	In combination with paclitaxel protein-bound for the treatment of adult patients with unresectable locally advanced or metastatic TNBC whose tumors express PD-L1 (PD-L1 stained tumor-infiltrating immune cells [IC] of any intensity covering ≥ 1% of the	Yes	paclitaxel protein-bound	Adult	unresectable locally advanced or metastatic	Not reported	Category 2A	Preferred

			tumor area), as determined by an FDA approved test.							
BAVENCIO	avelumab	Urothelial Carcinoma (UC)	Patients with locally advanced or metastatic urothelial carcinoma (UC) who have disease progression during or following platinum-containing chemotherapy	No	none	none	locally advanced or metastatic	Second-line	Category 2A	Alternative Preferred
BAVENCIO	avelumab	Urothelial Carcinoma (UC)	Patients with locally advanced or metastatic urothelial carcinoma (UC) who have disease progression within 12 months of neoadjuvant or adjuvant treatment with platinum-containing chemotherapy	No	none	none	locally advanced or metastatic	Second-line	Category 2A	Alternative Preferred
IMFINZI	durvalumab	Urothelial Carcinoma (UC)	Locally advanced or metastatic urothelial carcinoma who have disease progression during or following platinum-containing chemotherapy	No	none	none	locally advanced or metastatic	Second-line	Category 2A	Alternative Preferred
IMFINZI	durvalumab	Urothelial Carcinoma (UC)	Locally advanced or metastatic urothelial carcinoma who have disease progression within 12 months of neoadjuvant or adjuvant treatment with platinum-containing chemotherapy	No	none	none	locally advanced or metastatic	Second-line	Category 2A	Alternative Preferred

KEYTRUDA	pembrolizumab	Urothelial Carcinoma (UC)	For the treatment of patients with locally advanced or metastatic urothelial carcinoma who have disease progression during or following platinum-containing chemotherapy or within 12 months of neoadjuvant or adjuvant treatment with platinum-containing chemotherapy	No	during or following platinum-containing chemotherapy or within 12 months of neoadjuvant or adjuvant treatment with platinum-containing chemotherapy	none	locally advanced or metastatic	Second-line	Category 1	Preferred
KEYTRUDA	pembrolizumab	Urothelial Carcinoma (UC)	For the treatment of patients with locally advanced or metastatic urothelial carcinoma who are not eligible for cisplatin containing chemotherapy and whose tumors express PD-L1 [Combined Positive Score (CPS) ≥10] as determined by an FDA-approved test, or in patients who are not eligible for any platinum-containing chemotherapy regardless of PD-L1 status	Yes	not eligible for cisplatincontaining chemotherapy or not eligible for any platinum-containing chemotherapy	none	locally advanced or metastatic	First-line	Category 2A	Preferred

KEYTRUDA	pembrolizumab	Urothelial Carcinoma (UC)	For the treatment of patients with Bacillus Calmette-Guerin (BCG)-unresponsive, high-risk, non-muscle invasive bladder cancer (NMIBC) with carcinoma in situ (CIS) with or without papillary tumors who are ineligible for or have elected not to undergo cystectomy	No	ineligible for or have elected not to undergo cystectomy	none	Bacillus Calmette-Guerin (BCG)-unresponsive, high-risk, non-muscle invasive bladder cancer (NMIBC) with carcinoma in situ (CIS) with or without papillary tumors	Not reported	Category 2A	Not Reported
OPDIVO	nivolumab	Urothelial Carcinoma (UC)	Patients with locally advanced or metastatic urothelial carcinoma who have disease progression during or following platinum-containing chemotherapy	No	platinum-containing chemotherapy	none	locally advanced or metastatic	Second-line	Category 2A	Alternative Preferred
OPDIVO	nivolumab	Urothelial Carcinoma (UC)	Patients with locally advanced or metastatic urothelial carcinoma who have disease progression within 12 months of neoadjuvant or adjuvant treatment with platinum-containing chemotherapy.	No	platinum-containing chemotherapy.	none	locally advanced or metastatic	Second-line	Category 2A	Alternative Preferred

TECENTRIQ	atezolizumab	Urothelial Carcinoma (UC)	For the treatment of adult patients with locally advanced or metastatic urothelial carcinoma who are not eligible for cisplatin-containing chemotherapy and whose tumors express PD-L1 (PD-L1 stained tumorinfiltrating immune cells [IC] covering ≥ 5% of the tumor area), as determined by an FDA-approved test	Yes	none	Adult	locally advanced or metastatic	First-line	Category 2A	Preferred
TECENTRIQ	atezolizumab	Urothelial Carcinoma (UC)	For the treatment of adult patients with locally advanced or metastatic urothelial carcinoma who are not eligible for any platinum-containing chemotherapy regardless of PD-L1 status	No	none	Adult	locally advanced or metastatic	First-line	Category 2A	Preferred
TECENTRIQ	atezolizumab	Urothelial Carcinoma (UC)	For the treatment of adult patients with locally advanced or metastatic urothelial carcinoma who have disease progression during or following any platinum-containing chemotherapy, or within 12 months of neoadjuvant or adjuvant chemotherapy.	No	none	Adult	locally advanced or metastatic	Second-line	Category 2A	Alternative Preferred

Appendix B

Unique ID	1	**Study ID**	NCT02392172	**Assessor**	KK
Ref or Label	Barlesi	**Aim**	adhering to intervention (the 'per-protocol' effect)	**The effect of adhering to intervention…**	failures in implementing the intervention that could have affected the outcome; non-adherence to their assigned intervention by trial participants
Experimental	avelumab	**Comparator**	docetaxel	**Source**	Journal article(s) with results of the trial
Outcome	Overall Survival (TPS @ 1%)	**Results**	HR 0.9 95% CI 0.72-1.12	**Weight**	1

Domain	Signalling question	Response	Comments
Bias arising from the randomization process	1.1 Was the allocation sequence random?	Y	
	1.2 Was the allocation sequence concealed until participants were enrolled and assigned to interventions?	Y	
	1.3 Did baseline differences between intervention groups suggest a problem with the randomization process?	N	
	Risk of bias judgement	**Low**	
Bias due to deviations from intended interventions	2.1 Were participants aware of their assigned intervention during the trial?	Y	
	2.2 Were carers and people delivering the interventions aware of participants' assigned intervention during the trial?	Y	
	2.3. [If applicable:] If Y/PY/NI to 2.1 or 2.2: Were important non-protocol interventions balanced across intervention groups?	NA	
	2.4. [If applicable:] Were there failures in implementing the intervention that could have affected the outcome?	N	A small difference seen in subjects randomized & received a dose was observced between groups. Unlikely to impact study outcome.
	2.5. [If applicable:] Was there non-adherence to the assigned intervention regimen that could have affected participants' outcomes?	N	
	2.6. If N/PN/NI to 2.3, or Y/PY/NI to 2.4 or 2.5: Was an appropriate analysis used to estimate the effect of adhering to the intervention?	NA	
	Risk of bias judgement	**Some concerns**	
Bias due to missing outcome data	3.1 Were data for this outcome available for all, or nearly all, participants randomized?	Y	
	3.2 If N/PN/NI to 3.1: Is there evidence that result was not biased by missing outcome data?	NA	
	3.3 If N/PN to 3.2: Could missingness in the outcome depend on its true value?	NA	
	3.4 If Y/PY/NI to 3.3: Is it likely that missingness in the outcome depended on its true value?	NA	
	Risk of bias judgement	**Low**	
Bias in measurement of the outcome	4.1 Was the method of measuring the outcome inappropriate?	N	
	4.2 Could measurement or ascertainment of the outcome have differed between intervention groups?	N	
	4.3 Were outcome assessors aware of the intervention received by study participants?	Y	
	4.4 If Y/PY/NI to 4.3: Could assessment of the outcome have been influenced by knowledge of intervention received?	N	
	4.5 If Y/PY/NI to 4.4: Is it likely that assessment of the outcome was influenced by knowledge of intervention received?	NA	
	Risk of bias judgement	**Low**	
Bias in selection of the reported result	5.1 Were the data that produced this result analysed in accordance with a pre-specified analysis plan that was finalized before unblinded outcome data were available for analysis?	Y	
	5.2 ... multiple eligible outcome measurements (e.g. scales, definitions, time points) within the outcome domain?	N	PD-L1 measurements established at central laboratory
	5.3 ... multiple eligible analyses of the data?	N	
	Risk of bias judgement	**Low**	
Overall bias	**Risk of bias judgement**	**Some concerns**	

Unique ID	2	**Study ID**	NCT02256436	**Assessor**	KK
Ref or Label	Bullmunt, J.	**Aim**	adhering to intervention (the 'per-protocol' effect)	**The effect of adhering to intervention…**	failures in implementing the intervention that could have affected the outcome; non-adherence to their assigned intervention by trial participants
Experimental	pembrolizumab	**Comparator**	IC Chemo	**Source**	Journal article(s) with results of the trial
Outcome	OS @ 10% CPS	**Results**	HR 0.57 95%CI .37-.88	**Weight**	1

Domain	Signalling question	Response	Comments
Bias arising from the randomization process	1.1 Was the allocation sequence random?	Y	
	1.2 Was the allocation sequence concealed until participants were enrolled and assigned to interventions?	Y	
	1.3 Did baseline differences between intervention groups suggest a problem with the randomization process?	N	
	Risk of bias judgement	**Low**	
Bias due to deviations from intended interventions	2.1 Were participants aware of their assigned intervention during the trial?	Y	
	2.2 Were carers and people delivering the interventions aware of participants' assigned intervention during the trial?	Y	
	2.3. [If applicable:] If Y/PY/NI to 2.1 or 2.2: Were important non-protocol interventions balanced across intervention groups?	NA	
	2.4. [If applicable:] Were there failures in implementing the intervention that could have affected the outcome?	N	a single digit difference seen in subjects randomized & received a dose was observced between groups. Unlikely to impact study
	2.5. [If applicable:] Was there non-adherence to the assigned intervention regimen that could have affected participants' outcomes?	N	
	2.6. If N/PN/NI to 2.3, or Y/PY/NI to 2.4 or 2.5: Was an appropriate analysis used to estimate the effect of adhering to the intervention?	NA	
	Risk of bias judgement	**Low**	
Bias due to missing outcome data	3.1 Were data for this outcome available for all, or nearly all, participants randomized?	Y	
	3.2 If N/PN/NI to 3.1: Is there evidence that result was not biased by missing outcome data?	NA	
	3.3 If N/PN to 3.2: Could missingness in the outcome depend on its true value?	NA	
	3.4 If Y/PY/NI to 3.3: Is it likely that missingness in the outcome depended on its true value?	NA	
	Risk of bias judgement	**Low**	
Bias in measurement of the outcome	4.1 Was the method of measuring the outcome inappropriate?	N	
	4.2 Could measurement or ascertainment of the outcome have differed between intervention groups?	N	
	4.3 Were outcome assessors aware of the intervention received by study participants?	Y	
	4.4 If Y/PY/NI to 4.3: Could assessment of the outcome have been influenced by knowledge of intervention received?	N	
	4.5 If Y/PY/NI to 4.4: Is it likely that assessment of the outcome was influenced by knowledge of intervention received?	NA	
	Risk of bias judgement	**Low**	
Bias in selection of the reported result	5.1 Were the data that produced this result analysed in accordance with a pre-specified analysis plan that was finalized before unblinded outcome data were available for analysis?	Y	
	5.2 ... multiple eligible outcome measurements (e.g. scales, definitions, time points) within the outcome domain?	N	PD-L1 measurements established at central laboratory
	5.3 ... multiple eligible analyses of the data?	N	
	Risk of bias judgement	**Low**	
Overall bias	**Risk of bias judgement**	**Low**	

Unique ID	3	**Study ID**	NCT02613507	**Assessor**	KK
Ref or Label	Wu	**Aim**	adhering to intervention (the 'per-protocol' effect)	**The effect of adhering to intervention…**	failures in implementing the intervention that could have affected the outcome; non-adherence to their assigned intervention by trial participants
Experimental	nivolumab	**Comparator**	docetaxel	**Source**	Journal article(s) with results of the trial
Outcome	OS @ 1% PD-L1	**Results**	HR 0.62 95%CI 0.45-0.87	**Weight**	1

Domain	Signalling question	Response	Comments
Bias arising from the randomization process	1.1 Was the allocation sequence random?	Y	
	1.2 Was the allocation sequence concealed until participants were enrolled and assigned to interventions?	Y	
	1.3 Did baseline differences between intervention groups suggest a problem with the randomization process?	N	
	Risk of bias judgement	**Low**	
Bias due to deviations from intended interventions	2.1 Were participants aware of their assigned intervention during the trial?	Y	
	2.2 Were carers and people delivering the interventions aware of participants' assigned intervention during the trial?	Y	
	2.3. [If applicable:] If Y/PY/NI to 2.1 or 2.2: Were important non-protocol interventions balanced across intervention groups?	NA	
	2.4. [If applicable:] Were there failures in implementing the intervention that could have affected the outcome?	N	a single digit difference seen in subjects randomized & received a dose was observced between groups. Unlikely to impact study
	2.5. [If applicable:] Was there non-adherence to the assigned intervention regimen that could have affected participants' outcomes?	N	
	2.6. If N/PN/NI to 2.3, or Y/PY/NI to 2.4 or 2.5: Was an appropriate analysis used to estimate the effect of adhering to the intervention?	NA	
	Risk of bias judgement	**Low**	
Bias due to missing outcome data	3.1 Were data for this outcome available for all, or nearly all, participants randomized?	Y	
	3.2 If N/PN/NI to 3.1: Is there evidence that result was not biased by missing outcome data?	NA	
	3.3 If N/PN to 3.2: Could missingness in the outcome depend on its true value?	NA	
	3.4 If Y/PY/NI to 3.3: Is it likely that missingness in the outcome depended on its true value?	NA	
	Risk of bias judgement	**Low**	
Bias in measurement of the outcome	4.1 Was the method of measuring the outcome inappropriate?	N	
	4.2 Could measurement or ascertainment of the outcome have differed between intervention groups?	N	
	4.3 Were outcome assessors aware of the intervention received by study participants?	Y	
	4.4 If Y/PY/NI to 4.3: Could assessment of the outcome have been influenced by knowledge of intervention received?	N	
	4.5 If Y/PY/NI to 4.4: Is it likely that assessment of the outcome was influenced by knowledge of intervention received?	NA	
	Risk of bias judgement	**Low**	
Bias in selection of the reported result	5.1 Were the data that produced this result analysed in accordance with a pre-specified analysis plan that was finalized before unblinded outcome data were available for analysis?	Y	
	5.2 ... multiple eligible outcome measurements (e.g. scales, definitions, time points) within the outcome domain?	N	PD-L1 expression was prospectively assessed from archived tumor tissue or recent biopsy in a central laboratory before patient randomization.
	5.3 ... multiple eligible analyses of the data?	N	
	Risk of bias judgement	**Low**	
Overall bias	**Risk of bias judgement**	**Some concerns**	

Unique ID	4	**Study ID**	NCT02625623	**Assessor**	KK
Ref or Label	Bang	**Aim**	adhering to intervention (the 'per-protocol' effect)	**The effect of adhering to intervention…**	failures in implementing the intervention that could have affected the outcome; non-adherence to their assigned intervention by trial participants
Experimental	avelumab	**Comparator**	IC Chemo	**Source**	
Outcome	OS @ 1% PD-L1	**Results**	HR 0.94 95%CI 0.57-1.55	**Weight**	1

Domain	Signalling question	Response	Comments
Bias arising from the randomization process	1.1 Was the allocation sequence random?	Y	
	1.2 Was the allocation sequence concealed until participants were enrolled and assigned to interventions?	Y	
	1.3 Did baseline differences between intervention groups suggest a problem with the randomization process?	N	
	Risk of bias judgement	**Low**	
Bias due to deviations from intended interventions	2.1 Were participants aware of their assigned intervention during the trial?	Y	
	2.2 Were carers and people delivering the interventions aware of participants' assigned intervention during the trial?	Y	
	2.3. [If applicable:] If Y/PY/NI to 2.1 or 2.2: Were important non-protocol interventions balanced across intervention groups?	NA	
	2.4. [If applicable:] Were there failures in implementing the intervention that could have affected the outcome?	N	single digit difference observed in fall-out from randomization to dosing between groups. Unlikely to bias outcomes.
	2.5. [If applicable:] Was there non-adherence to the assigned intervention regimen that could have affected participants' outcomes?	N	
	2.6. If N/PN/NI to 2.3, or Y/PY/NI to 2.4 or 2.5: Was an appropriate analysis used to estimate the effect of adhering to the intervention?	NA	
	Risk of bias judgement	**Low**	
Bias due to missing outcome data	3.1 Were data for this outcome available for all, or nearly all, participants randomized?	Y	
	3.2 If N/PN/NI to 3.1: Is there evidence that result was not biased by missing outcome data?	NA	
	3.3 If N/PN to 3.2: Could missingness in the outcome depend on its true value?	NA	
	3.4 If Y/PY/NI to 3.3: Is it likely that missingness in the outcome depended on its true value?	NA	
	Risk of bias judgement	**Low**	
Bias in measurement of the outcome	4.1 Was the method of measuring the outcome inappropriate?	N	
	4.2 Could measurement or ascertainment of the outcome have differed between intervention groups?	N	
	4.3 Were outcome assessors aware of the intervention received by study participants?	Y	
	4.4 If Y/PY/NI to 4.3: Could assessment of the outcome have been influenced by knowledge of intervention received?	N	
	4.5 If Y/PY/NI to 4.4: Is it likely that assessment of the outcome was influenced by knowledge of intervention received?	NA	
	Risk of bias judgement	**Low**	
Bias in selection of the reported result	5.1 Were the data that produced this result analysed in accordance with a pre-specified analysis plan that was finalized before unblinded outcome data were available for analysis?	Y	
	5.2 ... multiple eligible outcome measurements (e.g. scales, definitions, time points) within the outcome domain?	N	PD-L1 retrospectively assessed at central laboratory
	5.3 ... multiple eligible analyses of the data?	N	
	Risk of bias judgement	**Low**	
Overall bias	**Risk of bias judgement**	**Low**	

Unique ID	5	**Study ID**	NCT02105636	**Assessor**	KK
Ref or Label	Ferris	**Aim**	adhering to intervention (the 'per-protocol' effect)	**The effect of adhering to intervention…**	failures in implementing the intervention that could have affected the outcome; non-adherence to their assigned intervention by trial participants
Experimental	nivolumab	**Comparator**	IC Chemo	**Source**	Journal article(s) with results of the trial
Outcome	OS	**Results**	HR 0.68 95%CI 0.54-0.86	**Weight**	1

Domain	Signalling question	Response	Comments
Bias arising from the randomization process	1.1 Was the allocation sequence random?	Y	
	1.2 Was the allocation sequence concealed until participants were enrolled and assigned to interventions?	Y	
	1.3 Did baseline differences between intervention groups suggest a problem with the randomization process?	N	
	Risk of bias judgement	**Low**	
Bias due to deviations from intended interventions	2.1 Were participants aware of their assigned intervention during the trial?	Y	
	2.2 Were carers and people delivering the interventions aware of participants' assigned intervention during the trial?	Y	
	2.3. [If applicable:] If Y/PY/NI to 2.1 or 2.2: Were important non-protocol interventions balanced across intervention groups?	NA	
	2.4. [If applicable:] Were there failures in implementing the intervention that could have affected the outcome?	N	
	2.5. [If applicable:] Was there non-adherence to the assigned intervention regimen that could have affected participants' outcomes?	N	
	2.6. If N/PN/NI to 2.3, or Y/PY/NI to 2.4 or 2.5: Was an appropriate analysis used to estimate the effect of adhering to the intervention?	NA	
	Risk of bias judgement	**Low**	
Bias due to missing outcome data	3.1 Were data for this outcome available for all, or nearly all, participants randomized?	Y	
	3.2 If N/PN/NI to 3.1: Is there evidence that result was not biased by missing outcome data?	NA	
	3.3 If N/PN to 3.2: Could missingness in the outcome depend on its true value?	NA	
	3.4 If Y/PY/NI to 3.3: Is it likely that missingness in the outcome depended on its true value?	NA	
	Risk of bias judgement	**Low**	
Bias in measurement of the outcome	4.1 Was the method of measuring the outcome inappropriate?	N	
	4.2 Could measurement or ascertainment of the outcome have differed between intervention groups?	N	
	4.3 Were outcome assessors aware of the intervention received by study participants?	Y	
	4.4 If Y/PY/NI to 4.3: Could assessment of the outcome have been influenced by knowledge of intervention received?	N	
	4.5 If Y/PY/NI to 4.4: Is it likely that assessment of the outcome was influenced by knowledge of intervention received?	NA	
	Risk of bias judgement	**Low**	
Bias in selection of the reported result	5.1 Were the data that produced this result analysed in accordance with a pre-specified analysis plan that was finalized before unblinded outcome data were available for analysis?	Y	
	5.2 ... multiple eligible outcome measurements (e.g. scales, definitions, time points) within the outcome domain?	N	Tumor PD-L1 membrane expression was evaluated centrally by means of IHC
	5.3 ... multiple eligible analyses of the data?	N	
	Risk of bias judgement	**Low**	Fresh or archived pretreatment tumor specimens were obtained after the last therapy and before trial entry from 90.6% of the patients.
Overall bias	**Risk of bias judgement**	**Low**	

Unique ID	6	**Study ID**	NCT01642004	**Assessor**	KK
Ref or Label	Brahmer	**Aim**	adhering to intervention (the 'per-protocol' effect)	**The effect of adhering to intervention…**	failures in implementing the intervention that could have affected the outcome; non-adherence to their assigned intervention by trial participants
Experimental	nivolumab	**Comparator**	docetaxel	**Source**	Journal article(s) with results of the trial
Outcome	OS @ 1%	**Results**	HR 0.69 95%CI 0.45-1.05	**Weight**	1

Domain	Signalling question	Response	Comments
Bias arising from the randomization process	1.1 Was the allocation sequence random?	Y	
	1.2 Was the allocation sequence concealed until participants were enrolled and assigned to interventions?	Y	
	1.3 Did baseline differences between intervention groups suggest a problem with the randomization process?	PN	patients were generally well□ balanced between the groups, with slight between-□
	Risk of bias judgement	**Low**	
Bias due to deviations from intended interventions	2.1 Were participants aware of their assigned intervention during the trial?	PY	
	2.2 Were carers and people delivering the interventions aware of participants' assigned intervention during the trial?	PY	
	2.3. [If applicable:] If Y/PY/NI to 2.1 or 2.2: Were important non-protocol interventions balanced across intervention groups?	NA	
	2.4. [If applicable:] Were there failures in implementing the intervention that could have affected the outcome?	N	
	2.5. [If applicable:] Was there non-adherence to the assigned intervention regimen that could have affected participants' outcomes?	N	
	2.6. If N/PN/NI to 2.3, or Y/PY/NI to 2.4 or 2.5: Was an appropriate analysis used to estimate the effect of adhering to the intervention?	NA	
	Risk of bias judgement	**Low**	A total of 83% of the patients who underwent randomization (225 of 272 patients) had quantifiable PD-L1 expression. Rates of PD-L1 positivity were balanced between the two treatment groups
Bias due to missing outcome data	3.1 Were data for this outcome available for all, or nearly all, participants randomized?	Y	
	3.2 If N/PN/NI to 3.1: Is there evidence that result was not biased by missing outcome data?	NA	
	3.3 If N/PN to 3.2: Could missingness in the outcome depend on its true value?	NA	
	3.4 If Y/PY/NI to 3.3: Is it likely that missingness in the outcome depended on its true value?	NA	
	Risk of bias judgement	**Low**	
Bias in measurement of the outcome	4.1 Was the method of measuring the outcome inappropriate?	N	
	4.2 Could measurement or ascertainment of the outcome have differed between intervention groups?	N	
	4.3 Were outcome assessors aware of the intervention received by study participants?	PY	
	4.4 If Y/PY/NI to 4.3: Could assessment of the outcome have been influenced by knowledge of intervention received?	N	
	4.5 If Y/PY/NI to 4.4: Is it likely that assessment of the outcome was influenced by knowledge of intervention received?	NA	
	Risk of bias judgement	**Low**	
Bias in selection of the reported result	5.1 Were the data that produced this result analysed in accordance with a pre-specified analysis plan that was finalized before unblinded outcome data were available for analysis?	Y	
	5.2 ... multiple eligible outcome measurements (e.g. scales, definitions, time points) within the outcome domain?	N	PD-L1 centrally assessed
	5.3 ... multiple eligible analyses of the data?	N	
	Risk of bias judgement	**Low**	
Overall bias	**Risk of bias judgement**	**Low**	

Unique ID	7	**Study ID**	NCT01668784	**Assessor**	KK
Ref or Label	Motzer	**Aim**	adhering to intervention (the 'per-protocol' effect)	**The effect of adhering to intervention…**	failures in implementing the intervention that could have affected the outcome; non-adherence to their assigned intervention by trial participants
Experimental	nivolumab	**Comparator**	everolimus	**Source**	Journal article(s) with results of the trial
Outcome	OS @ 1% PD-L1	**Results**	HR 0.79 95%CI 0.53-1.17	**Weight**	1

Domain	Signalling question	Response	Comments
Bias arising from the randomization process	1.1 Was the allocation sequence random?	Y	
	1.2 Was the allocation sequence concealed until participants were enrolled and assigned to interventions?	Y	
	1.3 Did baseline differences between intervention groups suggest a problem with the randomization process?	N	
	Risk of bias judgement	**Low**	92% of subjects had baseline tumor specimens available for PD-L1 testing.
Bias due to deviations from intended interventions	2.1 Were participants aware of their assigned intervention during the trial?	Y	
	2.2 Were carers and people delivering the interventions aware of participants' assigned intervention during the trial?	Y	
	2.3. [If applicable:] If Y/PY/NI to 2.1 or 2.2: Were important non-protocol interventions balanced across intervention groups?	NA	
	2.4. [If applicable:] Were there failures in implementing the intervention that could have affected the outcome?	N	
	2.5. [If applicable:] Was there non-adherence to the assigned intervention regimen that could have affected participants' outcomes?	N	
	2.6. If N/PN/NI to 2.3, or Y/PY/NI to 2.4 or 2.5: Was an appropriate analysis used to estimate the effect of adhering to the intervention?	NA	
	Risk of bias judgement	**Low**	
Bias due to missing outcome data	3.1 Were data for this outcome available for all, or nearly all, participants randomized?	Y	
	3.2 If N/PN/NI to 3.1: Is there evidence that result was not biased by missing outcome data?	NA	
	3.3 If N/PN to 3.2: Could missingness in the outcome depend on its true value?	NA	
	3.4 If Y/PY/NI to 3.3: Is it likely that missingness in the outcome depended on its true value?	NA	
	Risk of bias judgement	**Low**	
Bias in measurement of the outcome	4.1 Was the method of measuring the outcome inappropriate?	N	
	4.2 Could measurement or ascertainment of the outcome have differed between intervention groups?	N	
	4.3 Were outcome assessors aware of the intervention received by study participants?	Y	
	4.4 If Y/PY/NI to 4.3: Could assessment of the outcome have been influenced by knowledge of intervention received?	N	
	4.5 If Y/PY/NI to 4.4: Is it likely that assessment of the outcome was influenced by knowledge of intervention received?	NA	
	Risk of bias judgement	**Low**	
Bias in selection of the reported result	5.1 Were the data that produced this result analysed in accordance with a pre-specified analysis plan that was finalized before unblinded outcome data were available for analysis?	Y	
	5.2 ... multiple eligible outcome measurements (e.g. scales, definitions, time points) within the outcome domain?	N	PD-L1 measurements centrally completed
	5.3 ... multiple eligible analyses of the data?	N	
	Risk of bias judgement	**Low**	
Overall bias	**Risk of bias judgement**	**Low**	

Unique ID	8	**Study ID**	NCT02041533	**Assessor**	KK
Ref or Label	Carbone	**Aim**	adhering to intervention (the 'per-protocol' effect)	**The effect of adhering to intervention…**	failures in implementing the intervention that could have affected the outcome; non-adherence to their assigned intervention by trial participants
Experimental	nivolumab	**Comparator**	IC chemo	**Source**	Journal article(s) with results of the trial
Outcome	OS @ 1%	**Results**	HR 1.07 95%CI 0.86-1.33	**Weight**	1

Domain	Signalling question	Response	Comments
Bias arising from the randomization process	1.1 Was the allocation sequence random?	Y	
	1.2 Was the allocation sequence concealed until participants were enrolled and assigned to interventions?	Y	
	1.3 Did baseline differences between intervention groups suggest a problem with the randomization process?	PY	Among all the patients, the baseline characteristics were generally balanced between the treatment groups. However, in the nivolumab group, the percentage of women was lower than that in the chemotherapy group (32% vs. 45%), as was the percentage of patients with a PD-L1 expression level of 50% or more (32% vs. 47%); the percentage of patients with liver metastases was slightly higher in the nivolumab group (20% vs. 13%). In addition, patients in the nivolumab group had a greater tumor burden (on the basis of the median sum of target-lesion diameters) than those in the chemotherapy group.
	Risk of bias judgement	**Some concerns**	Only patients with a PD-L1 expression level of 1% or more underwent randomization. Doesn't include the negatives for comparison.
Bias due to deviations from intended interventions	2.1 Were participants aware of their assigned intervention during the trial?	Y	
	2.2 Were carers and people delivering the interventions aware of participants' assigned intervention during the trial?	Y	
	2.3. [If applicable:] If Y/PY/NI to 2.1 or 2.2: Were important non-protocol interventions balanced across intervention groups?	NA	
	2.4. [If applicable:] Were there failures in implementing the intervention that could have affected the outcome?	N	
	2.5. [If applicable:] Was there non-adherence to the assigned intervention regimen that could have affected participants' outcomes?	N	
	2.6. If N/PN/NI to 2.3, or Y/PY/NI to 2.4 or 2.5: Was an appropriate analysis used to estimate the effect of adhering to the intervention?	NA	
	Risk of bias judgement	**Low**	
Bias due to missing outcome data	3.1 Were data for this outcome available for all, or nearly all, participants randomized?	Y	
	3.2 If N/PN/NI to 3.1: Is there evidence that result was not biased by missing outcome data?	NA	
	3.3 If N/PN to 3.2: Could missingness in the outcome depend on its true value?	NA	
	3.4 If Y/PY/NI to 3.3: Is it likely that missingness in the outcome depended on its true value?	NA	
	Risk of bias judgement	**Low**	
Bias in measurement of the outcome	4.1 Was the method of measuring the outcome inappropriate?	N	
	4.2 Could measurement or ascertainment of the outcome have differed between intervention groups?	N	
	4.3 Were outcome assessors aware of the intervention received by study participants?	Y	
	4.4 If Y/PY/NI to 4.3: Could assessment of the outcome have been influenced by knowledge of intervention received?	N	
	4.5 If Y/PY/NI to 4.4: Is it likely that assessment of the outcome was influenced by knowledge of intervention received?	NA	
	Risk of bias judgement	**Low**	
Bias in selection of the reported result	5.1 Were the data that produced this result analysed in accordance with a pre-specified analysis plan that was finalized before unblinded outcome data were available for analysis?	Y	
	5.2 ... multiple eligible outcome measurements (e.g. scales, definitions, time points) within the outcome domain?	N	PD-L1 Centrally assessed
	5.3 ... multiple eligible analyses of the data?	N	
	Risk of bias judgement	**Low**	
Overall bias	**Risk of bias judgement**	**Some concerns**	Nivolumab group was biased toward more high PD-L1 expressors and more females. However, overall trial result was negative and did not find a difference between the 2 groups.

Unique ID	9	**Study ID**	NCT01721746	**Assessor**	KK
Ref or Label	Weber	**Aim**	adhering to intervention (the 'per-protocol' effect)	**The effect of adhering to intervention…**	failures in implementing the intervention that could have affected the outcome; non-adherence to their assigned intervention by trial participants
Experimental	nivolumab	**Comparator**	IC Chemo	**Source**	Journal article(s) with results of the trial
Outcome	OS @ 5% PD-L1	**Results**	HR 0.73 95%CI 0.49-1.09	**Weight**	1

Domain	Signalling question	Response	Comments
Bias arising from the randomization process	1.1 Was the allocation sequence random?	Y	
	1.2 Was the allocation sequence concealed until participants were enrolled and assigned to interventions?	Y	
	1.3 Did baseline differences between intervention groups suggest a problem with the randomization process?	N	Baseline characteristics were similar in the nivolumab and ICC study groups, with the exception of history of brain metastases and high lactate dehydrogenase, which were higher in the nivolumab than the ICC group. Additionally, patients were heavily treated, with half of all randomly allocated patients receiving at least two previous systemic therapies. The proportion of patients receiving previous immunotherapies other than ipilimumab was also higher in the ICC than the nivolumab group. The type and extent of previous treatments were generally consistent between treatment groups.
	Risk of bias judgement	**Low**	We stratified randomisation by tumour PD-L1 status by immunohistochemistry using an automated Bristol-Myers Squibb (New Jersey, USA)/Dako (California, USA) assay19 (positive in at least 5% of tumour cells vs negative or indeterminate)
Bias due to deviations from intended interventions	2.1 Were participants aware of their assigned intervention during the trial?	Y	
	2.2 Were carers and people delivering the interventions aware of participants' assigned intervention during the trial?	Y	
	2.3. [If applicable:] If Y/PY/NI to 2.1 or 2.2: Were important non-protocol interventions balanced across intervention groups?	NA	
	2.4. [If applicable:] Were there failures in implementing the intervention that could have affected the outcome?	N	
	2.5. [If applicable:] Was there non-adherence to the assigned intervention regimen that could have affected participants' outcomes?	N	4 vs 31 (nivo vs doce) did not recieve a dose, however unlikely to biat outcome.
	2.6. If N/PN/NI to 2.3, or Y/PY/NI to 2.4 or 2.5: Was an appropriate analysis used to estimate the effect of adhering to the intervention?	NA	
	Risk of bias judgement	**Low**	
Bias due to missing outcome data	3.1 Were data for this outcome available for all, or nearly all, participants randomized?	Y	
	3.2 If N/PN/NI to 3.1: Is there evidence that result was not biased by missing outcome data?	NA	
	3.3 If N/PN to 3.2: Could missingness in the outcome depend on its true value?	NA	
	3.4 If Y/PY/NI to 3.3: Is it likely that missingness in the outcome depended on its true value?	NA	
	Risk of bias judgement	**Low**	
Bias in measurement of the outcome	4.1 Was the method of measuring the outcome inappropriate?	N	
	4.2 Could measurement or ascertainment of the outcome have differed between intervention groups?	N	
	4.3 Were outcome assessors aware of the intervention received by study participants?	N	radiologists were blinded to treatment for tumor size measurements
	4.4 If Y/PY/NI to 4.3: Could assessment of the outcome have been influenced by knowledge of intervention received?	NA	
	4.5 If Y/PY/NI to 4.4: Is it likely that assessment of the outcome was influenced by knowledge of intervention received?	NA	
	Risk of bias judgement	**Low**	
Bias in selection of the reported result	5.1 Were the data that produced this result analysed in accordance with a pre-specified analysis plan that was finalized before unblinded outcome data were available for analysis?	Y	
	5.2 ... multiple eligible outcome measurements (e.g. scales, definitions, time points) within the outcome domain?	N	PD-L1 centrally assessed and assessors were unaware of group assignment
	5.3 ... multiple eligible analyses of the data?	N	
	Risk of bias judgement	**Low**	
Overall bias	**Risk of bias judgement**	**Low**	

Unique ID	10	**Study ID**	NCT01673867	**Assessor**	KK
Ref or Label	Borghei	**Aim**	adhering to intervention (the 'per-protocol' effect)	**The effect of adhering to intervention…**	failures in implementing the intervention that could have affected the outcome; non-adherence to their assigned intervention by trial participants
Experimental	nivolumab	**Comparator**	docetaxel	**Source**	Journal article(s) with results of the trial
Outcome	OS @ 1% PD-L1	**Results**	HR 0.58 95% CI 0.43-0.79	**Weight**	1

Domain	Signalling question	Response	Comments
Bias arising from the randomization process	1.1 Was the allocation sequence random?	Y	
	1.2 Was the allocation sequence concealed until participants were enrolled and assigned to interventions?	PY	
	1.3 Did baseline differences between intervention groups suggest a problem with the randomization process?	N	The baseline characteristics were balanced between the treatment groups, with slight between-group imbalances in the percentages of male patients and patients younger than 65 years of age.
	Risk of bias judgement	**Low**	
Bias due to deviations from intended interventions	2.1 Were participants aware of their assigned intervention during the trial?	Y	
	2.2 Were carers and people delivering the interventions aware of participants' assigned intervention during the trial?	Y	
	2.3. [If applicable:] If Y/PY/NI to 2.1 or 2.2: Were important non-protocol interventions balanced across intervention groups?	NA	
	2.4. [If applicable:] Were there failures in implementing the intervention that could have affected the outcome?	N	Five patients in the nivolumab group and 22 in the docetaxel group did not receive the assigned study drug
	2.5. [If applicable:] Was there non-adherence to the assigned intervention regimen that could have affected participants' outcomes?	N	
	2.6. If N/PN/NI to 2.3, or Y/PY/NI to 2.4 or 2.5: Was an appropriate analysis used to estimate the effect of adhering to the intervention?	NA	
	Risk of bias judgement	**Low**	Among the 582 patients who underwent randomization, 455 (78%) had quantifiable PD-L1 expression. Rates of PD-L1 expression were balanced between the two groups
Bias due to missing outcome data	3.1 Were data for this outcome available for all, or nearly all, participants randomized?	Y	
	3.2 If N/PN/NI to 3.1: Is there evidence that result was not biased by missing outcome data?	NA	
	3.3 If N/PN to 3.2: Could missingness in the outcome depend on its true value?	NA	
	3.4 If Y/PY/NI to 3.3: Is it likely that missingness in the outcome depended on its true value?	NA	
	Risk of bias judgement	**Low**	
Bias in measurement of the outcome	4.1 Was the method of measuring the outcome inappropriate?	N	
	4.2 Could measurement or ascertainment of the outcome have differed between intervention groups?	N	
	4.3 Were outcome assessors aware of the intervention received by study participants?	Y	
	4.4 If Y/PY/NI to 4.3: Could assessment of the outcome have been influenced by knowledge of intervention received?	PN	
	4.5 If Y/PY/NI to 4.4: Is it likely that assessment of the outcome was influenced by knowledge of intervention received?	NA	
	Risk of bias judgement	**Low**	
Bias in selection of the reported result	5.1 Were the data that produced this result analysed in accordance with a pre-specified analysis plan that was finalized before unblinded outcome data were available for analysis?	Y	
	5.2 ... multiple eligible outcome measurements (e.g. scales, definitions, time points) within the outcome domain?	N	Central lab assessment of PD-L1
	5.3 ... multiple eligible analyses of the data?	N	
	Risk of bias judgement	**Low**	
Overall bias	**Risk of bias judgement**	**Low**	

Unique ID	11	**Study ID**	NCT01721772	**Assessor**	KK
Ref or Label	Robert	**Aim**	adhering to intervention (the 'per-protocol' effect)	**The effect of adhering to intervention…**	failures in implementing the intervention that could have affected the outcome; non-adherence to their assigned intervention by trial participants
Experimental	nivolumab	**Comparator**	dacarbazine	**Source**	Journal article(s) with results of the trial
Outcome	OS @ 5% PD-L1	**Results**	HR 0.3 95%CI 0.15-0.60	**Weight**	1

Domain	Signalling question	Response	Comments
Bias arising from the randomization process	1.1 Was the allocation sequence random?	Y	
	1.2 Was the allocation sequence concealed until participants were enrolled and assigned to interventions?	Y	
	1.3 Did baseline differences between intervention groups suggest a problem with the randomization process?	N	
	Risk of bias judgement	**Low**	Prospective asseddment of PD-L1 @5% was used to stratify randomization
Bias due to deviations from intended interventions	2.1 Were participants aware of their assigned intervention during the trial?	Y	
	2.2 Were carers and people delivering the interventions aware of participants' assigned intervention during the trial?	Y	
	2.3. [If applicable:] If Y/PY/NI to 2.1 or 2.2: Were important non-protocol interventions balanced across intervention groups?	NA	
	2.4. [If applicable:] Were there failures in implementing the intervention that could have affected the outcome?	N	
	2.5. [If applicable:] Was there non-adherence to the assigned intervention regimen that could have affected participants' outcomes?	N	
	2.6. If N/PN/NI to 2.3, or Y/PY/NI to 2.4 or 2.5: Was an appropriate analysis used to estimate the effect of adhering to the intervention?	NA	
	Risk of bias judgement	**Low**	
Bias due to missing outcome data	3.1 Were data for this outcome available for all, or nearly all, participants randomized?	Y	
	3.2 If N/PN/NI to 3.1: Is there evidence that result was not biased by missing outcome data?	NA	
	3.3 If N/PN to 3.2: Could missingness in the outcome depend on its true value?	NA	
	3.4 If Y/PY/NI to 3.3: Is it likely that missingness in the outcome depended on its true value?	NA	
	Risk of bias judgement	**Low**	
Bias in measurement of the outcome	4.1 Was the method of measuring the outcome inappropriate?	N	
	4.2 Could measurement or ascertainment of the outcome have differed between intervention groups?	N	
	4.3 Were outcome assessors aware of the intervention received by study participants?	PY	
	4.4 If Y/PY/NI to 4.3: Could assessment of the outcome have been influenced by knowledge of intervention received?	N	
	4.5 If Y/PY/NI to 4.4: Is it likely that assessment of the outcome was influenced by knowledge of intervention received?	NA	
	Risk of bias judgement	**Low**	
Bias in selection of the reported result	5.1 Were the data that produced this result analysed in accordance with a pre-specified analysis plan that was finalized before unblinded outcome data were available for analysis?	Y	
	5.2 ... multiple eligible outcome measurements (e.g. scales, definitions, time points) within the outcome domain?	N	PD-L1 done centrally and prospectively
	5.3 ... multiple eligible analyses of the data?	N	
	Risk of bias judgement	**Low**	
Overall bias	**Risk of bias judgement**	**Low**	

Unique ID	12	**Study ID**	NCT01844505	**Assessor**	KK
Ref or Label	Larkin	**Aim**	adhering to intervention (the 'per-protocol' effect)	**The effect of adhering to intervention…**	failures in implementing the intervention that could have affected the outcome; non-adherence to their assigned intervention by trial participants
Experimental	nivolumab	**Comparator**	ipilmumab	**Source**	Journal article(s) with results of the trial
Outcome	OS @ 5% PD-L1	**Results**	HR NA Median NR 95% CI 39.1 - NR	**Weight**	1

Domain	Signalling question	Response	comments
Bias arising from the randomization process	1.1 Was the allocation sequence random?	Y	
	1.2 Was the allocation sequence concealed until participants were enrolled and assigned to interventions?	Y	
	1.3 Did baseline differences between intervention groups suggest a problem with the randomization process?	N	
	Risk of bias judgement	**Low**	
Bias due to deviations from intended interventions	2.1 Were participants aware of their assigned intervention during the trial?	N	
	2.2 Were carers and people delivering the interventions aware of participants' assigned intervention during the trial?	N	
	2.3. [If applicable:] If Y/PY/NI to 2.1 or 2.2: Were important non-protocol interventions balanced across intervention groups?	NA	
	2.4. [If applicable:] Were there failures in implementing the intervention that could have affected the outcome?	N	
	2.5. [If applicable:] Was there non-adherence to the assigned intervention regimen that could have affected participants' outcomes?	N	
	2.6. If N/PN/NI to 2.3, or Y/PY/NI to 2.4 or 2.5: Was an appropriate analysis used to estimate the effect of adhering to the intervention?	NA	
	Risk of bias judgement	**Low**	
Bias due to missing outcome data	3.1 Were data for this outcome available for all, or nearly all, participants randomized?	Y	
	3.2 If N/PN/NI to 3.1: Is there evidence that result was not biased by missing outcome data?	NA	
	3.3 If N/PN to 3.2: Could missingness in the outcome depend on its true value?	NA	
	3.4 If Y/PY/NI to 3.3: Is it likely that missingness in the outcome depended on its true value?	NA	
	Risk of bias judgement	**Low**	
Bias in measurement of the outcome	4.1 Was the method of measuring the outcome inappropriate?	N	
	4.2 Could measurement or ascertainment of the outcome have differed between intervention groups?	N	
	4.3 Were outcome assessors aware of the intervention received by study participants?	N	
	4.4 If Y/PY/NI to 4.3: Could assessment of the outcome have been influenced by knowledge of intervention received?	NA	
	4.5 If Y/PY/NI to 4.4: Is it likely that assessment of the outcome was influenced by knowledge of intervention received?	NA	
	Risk of bias judgement	**Low**	
Bias in selection of the reported result	5.1 Were the data that produced this result analysed in accordance with a pre-specified analysis plan that was finalized before unblinded outcome data were available for analysis?	Y	
	5.2 ... multiple eligible outcome measurements (e.g. scales, definitions, time points) within the outcome domain?	N	Prospective and centrally assessed PD-L1 measurements
	5.3 ... multiple eligible analyses of the data?	N	
	Risk of bias judgement	**Low**	
Overall bias	**Risk of bias judgement**	**Low**	

Unique ID	13	**Study ID**	NCT02788279	**Assessor**	KK
Ref or Label	Eng	**Aim**	adhering to intervention (the 'per-protocol' effect)	**The effect of adhering to intervention…**	failures in implementing the intervention that could have affected the outcome; non-adherence to their assigned intervention by trial participants
Experimental	atezolizumab	**Comparator**	regorafenib	**Source**	Journal article(s) with results of the trial
Outcome	ORR @5% PD-L1 TILS	**Results**	ORR 3%	**Weight**	1

Domain	Signalling question	Response	Comments
Bias arising from the randomization process	1.1 Was the allocation sequence random?	Y	
	1.2 Was the allocation sequence concealed until participants were enrolled and assigned to interventions?	Y	
	1.3 Did baseline differences between intervention groups suggest a problem with the randomization process?	N	
	Risk of bias judgement	**Low**	
Bias due to deviations from intended interventions	2.1 Were participants aware of their assigned intervention during the trial?	Y	
	2.2 Were carers and people delivering the interventions aware of participants' assigned intervention during the trial?	Y	
	2.3. [If applicable:] If Y/PY/NI to 2.1 or 2.2: Were important non-protocol interventions balanced across intervention groups?	NA	
	2.4. [If applicable:] Were there failures in implementing the intervention that could have affected the outcome?	N	single digit differences between groups in who received treatment post randomization.
	2.5. [If applicable:] Was there non-adherence to the assigned intervention regimen that could have affected participants' outcomes?	N	
	2.6. If N/PN/NI to 2.3, or Y/PY/NI to 2.4 or 2.5: Was an appropriate analysis used to estimate the effect of adhering to the intervention?	NA	
	Risk of bias judgement	**Low**	
Bias due to missing outcome data	3.1 Were data for this outcome available for all, or nearly all, participants randomized?	Y	
	3.2 If N/PN/NI to 3.1: Is there evidence that result was not biased by missing outcome data?	NA	
	3.3 If N/PN to 3.2: Could missingness in the outcome depend on its true value?	NA	
	3.4 If Y/PY/NI to 3.3: Is it likely that missingness in the outcome depended on its true value?	NA	
	Risk of bias judgement	**Low**	
Bias in measurement of the outcome	4.1 Was the method of measuring the outcome inappropriate?	N	
	4.2 Could measurement or ascertainment of the outcome have differed between intervention groups?	N	
	4.3 Were outcome assessors aware of the intervention received by study participants?	PY	
	4.4 If Y/PY/NI to 4.3: Could assessment of the outcome have been influenced by knowledge of intervention received?	N	
	4.5 If Y/PY/NI to 4.4: Is it likely that assessment of the outcome was influenced by knowledge of intervention received?	NA	
	Risk of bias judgement	**Low**	
Bias in selection of the reported result	5.1 Were the data that produced this result analysed in accordance with a pre-specified analysis plan that was finalized before unblinded outcome data were available for analysis?	Y	
	5.2 ... multiple eligible outcome measurements (e.g. scales, definitions, time points) within the outcome domain?	N	central analysis of PD-L1
	5.3 ... multiple eligible analyses of the data?	N	
	Risk of bias judgement	**Low**	
Overall bias	**Risk of bias judgement**	**Some concerns**	Complete analysis not published due to negative trial result

Unique ID	14	**Study ID**	NCT02302807	**Assessor**	KK
Ref or Label	Powles	**Aim**	adhering to intervention (the 'per-protocol' effect)	**The effect of adhering to intervention…**	failures in implementing the intervention that could have affected the outcome; non-adherence to their assigned intervention by trial participants
Experimental	atezolizumab	**Comparator**	IC Chemo	**Source**	Journal article(s) with results of the trial
Outcome	OS @ 5% PD-L1 IC2/3	**Results**	HR .87 95%CI 0.63-1.21	**Weight**	1

Domain	Signalling question	Response	Comments
Bias arising from the randomization process	1.1 Was the allocation sequence random?	Y	
	1.2 Was the allocation sequence concealed until participants were enrolled and assigned to interventions?	Y	
	1.3 Did baseline differences between intervention groups suggest a problem with the randomization process?	N	
	Risk of bias judgement	**Low**	PD-L1 status was blinded to all participants throughout trial duration
Bias due to deviations from intended interventions	2.1 Were participants aware of their assigned intervention during the trial?	Y	
	2.2 Were carers and people delivering the interventions aware of participants' assigned intervention during the trial?	Y	
	2.3. [If applicable:] If Y/PY/NI to 2.1 or 2.2: Were important non-protocol interventions balanced across intervention groups?	NA	
	2.4. [If applicable:] Were there failures in implementing the intervention that could have affected the outcome?	N	small numerical differences between groups regarding who declined participation after randomization
	2.5. [If applicable:] Was there non-adherence to the assigned intervention regimen that could have affected participants' outcomes?	N	
	2.6. If N/PN/NI to 2.3, or Y/PY/NI to 2.4 or 2.5: Was an appropriate analysis used to estimate the effect of adhering to the intervention?	NA	
	Risk of bias judgement	**Low**	
Bias due to missing outcome data	3.1 Were data for this outcome available for all, or nearly all, participants randomized?	Y	
	3.2 If N/PN/NI to 3.1: Is there evidence that result was not biased by missing outcome data?	NA	
	3.3 If N/PN to 3.2: Could missingness in the outcome depend on its true value?	NA	
	3.4 If Y/PY/NI to 3.3: Is it likely that missingness in the outcome depended on its true value?	NA	
	Risk of bias judgement	**Low**	
Bias in measurement of the outcome	4.1 Was the method of measuring the outcome inappropriate?	N	
	4.2 Could measurement or ascertainment of the outcome have differed between intervention groups?	N	
	4.3 Were outcome assessors aware of the intervention received by study participants?	PY	
	4.4 If Y/PY/NI to 4.3: Could assessment of the outcome have been influenced by knowledge of intervention received?	N	
	4.5 If Y/PY/NI to 4.4: Is it likely that assessment of the outcome was influenced by knowledge of intervention received?	NA	
	Risk of bias judgement	**Low**	
Bias in selection of the reported result	5.1 Were the data that produced this result analysed in accordance with a pre-specified analysis plan that was finalized before unblinded outcome data were available for analysis?	Y	
	5.2 ... multiple eligible outcome measurements (e.g. scales, definitions, time points) within the outcome domain?	N	PD-L1 assessed centrally
	5.3 ... multiple eligible analyses of the data?	N	
	Risk of bias judgement	**Low**	
Overall bias	**Risk of bias judgement**	**Low**	

Unique ID	15	**Study ID**	NCT01704287	**Assessor**	KK
Ref or Label	Hamid	**Aim**	adhering to intervention (the 'per-protocol' effect)	**The effect of adhering to intervention…**	failures in implementing the intervention that could have affected the outcome; non-adherence to their assigned intervention by trial participants
Experimental	pembrolizumab	**Comparator**	IC Chemo	**Source**	Journal article(s) with results of the trial
Outcome	PFS @ 1% PD-L1	**Results**	HR 0.52 95%CI 0.39-0.68	**Weight**	1

Domain	Signalling question	Response	Comments
Bias arising from the randomization process	1.1 Was the allocation sequence random?	Y	
	1.2 Was the allocation sequence concealed until participants were enrolled and assigned to interventions?	Y	
	1.3 Did baseline differences between intervention groups suggest a problem with the randomization process?	N	
	Risk of bias judgement	**Low**	
Bias due to deviations from intended interventions	2.1 Were participants aware of their assigned intervention during the trial?	N	
	2.2 Were carers and people delivering the interventions aware of participants' assigned intervention during the trial?	N	
	2.3. [If applicable:] If Y/PY/NI to 2.1 or 2.2: Were important non-protocol interventions balanced across intervention groups?	NA	
	2.4. [If applicable:] Were there failures in implementing the intervention that could have affected the outcome?	N	
	2.5. [If applicable:] Was there non-adherence to the assigned intervention regimen that could have affected participants' outcomes?	N	
	2.6. If N/PN/NI to 2.3, or Y/PY/NI to 2.4 or 2.5: Was an appropriate analysis used to estimate the effect of adhering to the intervention?	NA	
	Risk of bias judgement	**Low**	
Bias due to missing outcome data	3.1 Were data for this outcome available for all, or nearly all, participants randomized?	Y	
	3.2 If N/PN/NI to 3.1: Is there evidence that result was not biased by missing outcome data?	NA	
	3.3 If N/PN to 3.2: Could missingness in the outcome depend on its true value?	NA	
	3.4 If Y/PY/NI to 3.3: Is it likely that missingness in the outcome depended on its true value?	NA	
	Risk of bias judgement	**Low**	
Bias in measurement of the outcome	4.1 Was the method of measuring the outcome inappropriate?	N	
	4.2 Could measurement or ascertainment of the outcome have differed between intervention groups?	N	
	4.3 Were outcome assessors aware of the intervention received by study participants?	N	
	4.4 If Y/PY/NI to 4.3: Could assessment of the outcome have been influenced by knowledge of intervention received?	NA	
	4.5 If Y/PY/NI to 4.4: Is it likely that assessment of the outcome was influenced by knowledge of intervention received?	NA	
	Risk of bias judgement	**Low**	
Bias in selection of the reported result	5.1 Were the data that produced this result analysed in accordance with a pre-specified analysis plan that was finalized before unblinded outcome data were available for analysis?	Y	
	5.2 ... multiple eligible outcome measurements (e.g. scales, definitions, time points) within the outcome domain?	N	
	5.3 ... multiple eligible analyses of the data?	N	
	Risk of bias judgement	**Low**	
Overall bias	**Risk of bias judgement**	**Low**	

Unique ID	16	**Study ID**	NCT01866319	**Assessor**	KK
Ref or Label	Robert	**Aim**	adhering to intervention (the 'per-protocol' effect)	**The effect of adhering to intervention…**	failures in implementing the intervention that could have affected the outcome; non-adherence to their assigned intervention by trial participants
Experimental	pembrolizumab	**Comparator**	ipilimumab	**Source**	Journal article(s) with results of the trial
Outcome	OS @ 1% PD-L1	**Results**	Median 35.6 95%CI 29.3-NR	**Weight**	1

Domain	Signalling question	Response	Comments
Bias arising from the randomization process	1.1 Was the allocation sequence random?	Y	
	1.2 Was the allocation sequence concealed until participants were enrolled and assigned to interventions?	Y	
	1.3 Did baseline differences between intervention groups suggest a problem with the randomization process?	N	
	Risk of bias judgement	**Low**	
Bias due to deviations from intended interventions	2.1 Were participants aware of their assigned intervention during the trial?	PY	
	2.2 Were carers and people delivering the interventions aware of participants' assigned intervention during the trial?	PY	
	2.3. [If applicable:] If Y/PY/NI to 2.1 or 2.2: Were important non-protocol interventions balanced across intervention groups?	NA	
	2.4. [If applicable:] Were there failures in implementing the intervention that could have affected the outcome?	PN	
	2.5. [If applicable:] Was there non-adherence to the assigned intervention regimen that could have affected participants' outcomes?	PN	
	2.6. If N/PN/NI to 2.3, or Y/PY/NI to 2.4 or 2.5: Was an appropriate analysis used to estimate the effect of adhering to the intervention?	NA	
	Risk of bias judgement	**Low**	
Bias due to missing outcome data	3.1 Were data for this outcome available for all, or nearly all, participants randomized?	Y	
	3.2 If N/PN/NI to 3.1: Is there evidence that result was not biased by missing outcome data?	NA	
	3.3 If N/PN to 3.2: Could missingness in the outcome depend on its true value?	NA	
	3.4 If Y/PY/NI to 3.3: Is it likely that missingness in the outcome depended on its true value?	NA	
	Risk of bias judgement	**Low**	
Bias in measurement of the outcome	4.1 Was the method of measuring the outcome inappropriate?	N	
	4.2 Could measurement or ascertainment of the outcome have differed between intervention groups?	N	
	4.3 Were outcome assessors aware of the intervention received by study participants?	PY	
	4.4 If Y/PY/NI to 4.3: Could assessment of the outcome have been influenced by knowledge of intervention received?	N	
	4.5 If Y/PY/NI to 4.4: Is it likely that assessment of the outcome was influenced by knowledge of intervention received?	NA	
	Risk of bias judgement	**Low**	
Bias in selection of the reported result	5.1 Were the data that produced this result analysed in accordance with a pre-specified analysis plan that was finalized before unblinded outcome data were available for analysis?	Y	
	5.2 ... multiple eligible outcome measurements (e.g. scales, definitions, time points) within the outcome domain?	N	
	5.3 ... multiple eligible analyses of the data?	N	
	Risk of bias judgement	**Low**	
Overall bias	**Risk of bias judgement**	**Low**	

Unique ID	17	**Study ID**	NCT01905657	**Assessor**	KK
Ref or Label	Herbst	**Aim**	adhering to intervention (the 'per-protocol' effect)	**The effect of adhering to intervention…**	failures in implementing the intervention that could have affected the outcome; non-adherence to their assigned intervention by trial participants
Experimental	pembrolizumab	**Comparator**	docetaxel	**Source**	Journal article(s) with results of the trial
Outcome	OS @ 1% PD-L1	**Results**	HR 0.66 95%CI 0.57-0.77	**Weight**	1

Domain	Signalling question	Response	Comments
Bias arising from the randomization process	1.1 Was the allocation sequence random?	Y	
	1.2 Was the allocation sequence concealed until participants were enrolled and assigned to interventions?	Y	
	1.3 Did baseline differences between intervention groups suggest a problem with the randomization process?	N	
	Risk of bias judgement	**Low**	
Bias due to deviations from intended interventions	2.1 Were participants aware of their assigned intervention during the trial?	Y	
	2.2 Were carers and people delivering the interventions aware of participants' assigned intervention during the trial?	Y	
	2.3. [If applicable:] If Y/PY/NI to 2.1 or 2.2: Were important non-protocol interventions balanced across intervention groups?	NA	
	2.4. [If applicable:] Were there failures in implementing the intervention that could have affected the outcome?	PN	Some numerical differences in dropouts from randomization to first dose between groups
	2.5. [If applicable:] Was there non-adherence to the assigned intervention regimen that could have affected participants' outcomes?	N	
	2.6. If N/PN/NI to 2.3, or Y/PY/NI to 2.4 or 2.5: Was an appropriate analysis used to estimate the effect of adhering to the intervention?	NA	
	Risk of bias judgement	**Low**	
Bias due to missing outcome data	3.1 Were data for this outcome available for all, or nearly all, participants randomized?	Y	
	3.2 If N/PN/NI to 3.1: Is there evidence that result was not biased by missing outcome data?	NA	
	3.3 If N/PN to 3.2: Could missingness in the outcome depend on its true value?	NA	
	3.4 If Y/PY/NI to 3.3: Is it likely that missingness in the outcome depended on its true value?	NA	
	Risk of bias judgement	**Low**	
Bias in measurement of the outcome	4.1 Was the method of measuring the outcome inappropriate?	N	
	4.2 Could measurement or ascertainment of the outcome have differed between intervention groups?	N	
	4.3 Were outcome assessors aware of the intervention received by study participants?	Y	
	4.4 If Y/PY/NI to 4.3: Could assessment of the outcome have been influenced by knowledge of intervention received?	N	
	4.5 If Y/PY/NI to 4.4: Is it likely that assessment of the outcome was influenced by knowledge of intervention received?	NA	
	Risk of bias judgement	**Low**	
Bias in selection of the reported result	5.1 Were the data that produced this result analysed in accordance with a pre-specified analysis plan that was finalized before unblinded outcome data were available for analysis?	Y	
	5.2 ... multiple eligible outcome measurements (e.g. scales, definitions, time points) within the outcome domain?	N	PD-L1 assessed centrally
	5.3 ... multiple eligible analyses of the data?	N	
	Risk of bias judgement	**Low**	
Overall bias	**Risk of bias judgement**	**Some concerns**	

Unique ID	18	**Study ID**	NCT02142738	**Assessor**	KK
Ref or Label	Reck	**Aim**	adhering to intervention (the 'per-protocol' effect)	**The effect of adhering to intervention…**	failures in implementing the intervention that could have affected the outcome; non-adherence to their assigned intervention by trial participants
Experimental	pembrolizumab	**Comparator**	IC Chemo	**Source**	Journal article(s) with results of the trial
Outcome	OS @ 50% PD-L1	**Results**	HR 0.63 95% CI 0.47-0.86	**Weight**	1

Domain	Signalling question	Response	Comments
Bias arising from the randomization process	1.1 Was the allocation sequence random?	Y	
	1.2 Was the allocation sequence concealed until participants were enrolled and assigned to interventions?	Y	
	1.3 Did baseline differences between intervention groups suggest a problem with the randomization process?	PN	Generally similar although more patients in the chemotherapy group than in the pembrolizumab group had never smoked (12.6% vs. 3.2%) and more patients in the pembrolizumab group than in the chemotherapy group had brain metastases (11.7% vs. 6.6%). These differences were not statistically significant.
	Risk of bias judgement	**Low**	
Bias due to deviations from intended interventions	2.1 Were participants aware of their assigned intervention during the trial?	Y	
	2.2 Were carers and people delivering the interventions aware of participants' assigned intervention during the trial?	Y	
	2.3. [If applicable:] If Y/PY/NI to 2.1 or 2.2: Were important non-protocol interventions balanced across intervention groups?	NA	
	2.4. [If applicable:] Were there failures in implementing the intervention that could have affected the outcome?	PN	
	2.5. [If applicable:] Was there non-adherence to the assigned intervention regimen that could have affected participants' outcomes?	PN	
	2.6. If N/PN/NI to 2.3, or Y/PY/NI to 2.4 or 2.5: Was an appropriate analysis used to estimate the effect of adhering to the intervention?	NA	
	Risk of bias judgement	**Low**	
Bias due to missing outcome data	3.1 Were data for this outcome available for all, or nearly all, participants randomized?	Y	
	3.2 If N/PN/NI to 3.1: Is there evidence that result was not biased by missing outcome data?	NA	
	3.3 If N/PN to 3.2: Could missingness in the outcome depend on its true value?	NA	
	3.4 If Y/PY/NI to 3.3: Is it likely that missingness in the outcome depended on its true value?	NA	
	Risk of bias judgement	**Low**	
Bias in measurement of the outcome	4.1 Was the method of measuring the outcome inappropriate?	N	
	4.2 Could measurement or ascertainment of the outcome have differed between intervention groups?	N	
	4.3 Were outcome assessors aware of the intervention received by study participants?	N	
	4.4 If Y/PY/NI to 4.3: Could assessment of the outcome have been influenced by knowledge of intervention received?	NA	
	4.5 If Y/PY/NI to 4.4: Is it likely that assessment of the outcome was influenced by knowledge of intervention received?	NA	
	Risk of bias judgement	**Low**	
Bias in selection of the reported result	5.1 Were the data that produced this result analysed in accordance with a pre-specified analysis plan that was finalized before unblinded outcome data were available for analysis?	Y	
	5.2 ... multiple eligible outcome measurements (e.g. scales, definitions, time points) within the outcome domain?	N	PD-L1 centrally assessed
	5.3 ... multiple eligible analyses of the data?	N	
	Risk of bias judgement	**Low**	
Overall bias	**Risk of bias judgement**	**Low**	

Unique ID	19	**Study ID**	NCT02252042	**Assessor**	KK
Ref or Label	Cohen, E	**Aim**	adhering to intervention (the 'per-protocol' effect)	**The effect of adhering to intervention…**	failures in implementing the intervention that could have affected the outcome; non-adherence to their assigned intervention by trial participants
Experimental	pembrolizumab	**Comparator**	IC Chemo	**Source**	Journal article(s) with results of the trial
Outcome	OS @50% PD-L1	**Results**	HR 0.53 95%CI 0.35-0.81	**Weight**	1

Domain	Signalling question	Response	Comments
Bias arising from the randomization process	1.1 Was the allocation sequence random?	Y	
	1.2 Was the allocation sequence concealed until participants were enrolled and assigned to interventions?	Y	
	1.3 Did baseline differences between intervention groups suggest a problem with the randomization process?	N	
	Risk of bias judgement	**Low**	
Bias due to deviations from intended interventions	2.1 Were participants aware of their assigned intervention during the trial?	Y	
	2.2 Were carers and people delivering the interventions aware of participants' assigned intervention during the trial?	Y	
	2.3. [If applicable:] If Y/PY/NI to 2.1 or 2.2: Were important non-protocol interventions balanced across intervention groups?	NA	
	2.4. [If applicable:] Were there failures in implementing the intervention that could have affected the outcome?	PN	10 on Chemo arm withdrew consent prior to dosing versus 1 in pembro group. Unlikely to impact outcomes
	2.5. [If applicable:] Was there non-adherence to the assigned intervention regimen that could have affected participants' outcomes?	N	
	2.6. If N/PN/NI to 2.3, or Y/PY/NI to 2.4 or 2.5: Was an appropriate analysis used to estimate the effect of adhering to the intervention?	NA	
	Risk of bias judgement	**Low**	
Bias due to missing outcome data	3.1 Were data for this outcome available for all, or nearly all, participants randomized?	Y	
	3.2 If N/PN/NI to 3.1: Is there evidence that result was not biased by missing outcome data?	NA	
	3.3 If N/PN to 3.2: Could missingness in the outcome depend on its true value?	NA	
	3.4 If Y/PY/NI to 3.3: Is it likely that missingness in the outcome depended on its true value?	NA	
	Risk of bias judgement	**Low**	
Bias in measurement of the outcome	4.1 Was the method of measuring the outcome inappropriate?	N	
	4.2 Could measurement or ascertainment of the outcome have differed between intervention groups?	N	
	4.3 Were outcome assessors aware of the intervention received by study participants?	N	
	4.4 If Y/PY/NI to 4.3: Could assessment of the outcome have been influenced by knowledge of intervention received?	NA	
	4.5 If Y/PY/NI to 4.4: Is it likely that assessment of the outcome was influenced by knowledge of intervention received?	NA	
	Risk of bias judgement	**Low**	
Bias in selection of the reported result	5.1 Were the data that produced this result analysed in accordance with a pre-specified analysis plan that was finalized before unblinded outcome data were available for analysis?	Y	
	5.2 ... multiple eligible outcome measurements (e.g. scales, definitions, time points) within the outcome domain?	N	
	5.3 ... multiple eligible analyses of the data?	N	
	Risk of bias judgement	**Low**	
Overall bias	**Risk of bias judgement**	**Low**	

Unique ID	20	**Study ID**	NCT02370498	**Assessor**	KK
Ref or Label	Shitara	**Aim**	adhering to intervention (the 'per-protocol' effect)	**The effect of adhering to intervention…**	failures in implementing the intervention that could have affected the outcome; non-adherence to their assigned intervention by trial participants
Experimental	pembrolizumab	**Comparator**	paclitaxel	**Source**	Journal article(s) with results of the trial
Outcome	OS @ 1CPS	**Results**	HR 0.82 95%CI 0.66-1.03	**Weight**	1

Domain	Signalling question	Response	Comments
Bias arising from the randomization process	1.1 Was the allocation sequence random?	Y	
	1.2 Was the allocation sequence concealed until participants were enrolled and assigned to interventions?	Y	
	1.3 Did baseline differences between intervention groups suggest a problem with the randomization process?	PN	more patients in paclitaxel group were >65. Some differences in histological subtypes
	Risk of bias judgement	**Low**	
Bias due to deviations from intended interventions	2.1 Were participants aware of their assigned intervention during the trial?	Y	
	2.2 Were carers and people delivering the interventions aware of participants' assigned intervention during the trial?	Y	
	2.3. [If applicable:] If Y/PY/NI to 2.1 or 2.2: Were important non-protocol interventions balanced across intervention groups?	NA	
	2.4. [If applicable:] Were there failures in implementing the intervention that could have affected the outcome?	PN	20 patients dropped out of paclitaxel group prior to dosing versus 2 in pembro group
	2.5. [If applicable:] Was there non-adherence to the assigned intervention regimen that could have affected participants' outcomes?	N	
	2.6. If N/PN/NI to 2.3, or Y/PY/NI to 2.4 or 2.5: Was an appropriate analysis used to estimate the effect of adhering to the intervention?	NA	
	Risk of bias judgement	**Low**	
Bias due to missing outcome data	3.1 Were data for this outcome available for all, or nearly all, participants randomized?	Y	
	3.2 If N/PN/NI to 3.1: Is there evidence that result was not biased by missing outcome data?	NA	
	3.3 If N/PN to 3.2: Could missingness in the outcome depend on its true value?	NA	
	3.4 If Y/PY/NI to 3.3: Is it likely that missingness in the outcome depended on its true value?	NA	
	Risk of bias judgement	**Low**	
Bias in measurement of the outcome	4.1 Was the method of measuring the outcome inappropriate?	N	
	4.2 Could measurement or ascertainment of the outcome have differed between intervention groups?	N	
	4.3 Were outcome assessors aware of the intervention received by study participants?	N	
	4.4 If Y/PY/NI to 4.3: Could assessment of the outcome have been influenced by knowledge of intervention received?	NA	
	4.5 If Y/PY/NI to 4.4: Is it likely that assessment of the outcome was influenced by knowledge of intervention received?	NA	
	Risk of bias judgement	**Low**	
Bias in selection of the reported result	5.1 Were the data that produced this result analysed in accordance with a pre-specified analysis plan that was finalized before unblinded outcome data were available for analysis?	Y	
	5.2 ... multiple eligible outcome measurements (e.g. scales, definitions, time points) within the outcome domain?	N	
	5.3 ... multiple eligible analyses of the data?	N	
	Risk of bias judgement	**Low**	
Overall bias	**Risk of bias judgement**	**Low**	

Unique ID	21	**Study ID**	NCT02220894	**Assessor**	KK
Ref or Label	Mok	**Aim**	adhering to intervention (the 'per-protocol' effect)	**The effect of adhering to intervention…**	failures in implementing the intervention that could have affected the outcome; non-adherence to their assigned intervention by trial participants
Experimental	pembrolizumab	**Comparator**	IC Chemo	**Source**	Journal article(s) with results of the trial
Outcome	OS @ 1% TPS	**Results**	HR 0.81 95%CI 0.71-0.93	**Weight**	1

Domain	Signalling question	Response	Comments
Bias arising from the randomization process	1.1 Was the allocation sequence random?	Y	
	1.2 Was the allocation sequence concealed until participants were enrolled and assigned to interventions?	Y	
	1.3 Did baseline differences between intervention groups suggest a problem with the randomization process?	N	
	Risk of bias judgement	**Low**	
Bias due to deviations from intended interventions	2.1 Were participants aware of their assigned intervention during the trial?	Y	
	2.2 Were carers and people delivering the interventions aware of participants' assigned intervention during the trial?	Y	
	2.3. [If applicable:] If Y/PY/NI to 2.1 or 2.2: Were important non-protocol interventions balanced across intervention groups?	NA	
	2.4. [If applicable:] Were there failures in implementing the intervention that could have affected the outcome?	PN	Over 60 subjects in chemo group withdrew consent prior to dosing as compared to zero in pembro group
	2.5. [If applicable:] Was there non-adherence to the assigned intervention regimen that could have affected participants' outcomes?	N	
	2.6. If N/PN/NI to 2.3, or Y/PY/NI to 2.4 or 2.5: Was an appropriate analysis used to estimate the effect of adhering to the intervention?	NA	
	Risk of bias judgement	**Low**	
Bias due to missing outcome data	3.1 Were data for this outcome available for all, or nearly all, participants randomized?	Y	
	3.2 If N/PN/NI to 3.1: Is there evidence that result was not biased by missing outcome data?	NA	
	3.3 If N/PN to 3.2: Could missingness in the outcome depend on its true value?	NA	
	3.4 If Y/PY/NI to 3.3: Is it likely that missingness in the outcome depended on its true value?	NA	
	Risk of bias judgement	**Low**	
Bias in measurement of the outcome	4.1 Was the method of measuring the outcome inappropriate?	N	
	4.2 Could measurement or ascertainment of the outcome have differed between intervention groups?	N	
	4.3 Were outcome assessors aware of the intervention received by study participants?	N	
	4.4 If Y/PY/NI to 4.3: Could assessment of the outcome have been influenced by knowledge of intervention received?	NA	
	4.5 If Y/PY/NI to 4.4: Is it likely that assessment of the outcome was influenced by knowledge of intervention received?	NA	
	Risk of bias judgement	**Low**	
Bias in selection of the reported result	5.1 Were the data that produced this result analysed in accordance with a pre-specified analysis plan that was finalized before unblinded outcome data were available for analysis?	Y	
	5.2 ... multiple eligible outcome measurements (e.g. scales, definitions, time points) within the outcome domain?	N	
	5.3 ... multiple eligible analyses of the data?	N	
	Risk of bias judgement	**Low**	
Overall bias	**Risk of bias judgement**	**Some concerns**	Inclusion Bias - Only included 1% positive cases, so no comparison to negatives is possible

Unique ID	22	**Study ID**	NCT02008227	**Assessor**	KK
Ref or Label	Fehrenbacher	**Aim**	adhering to intervention (the 'per-protocol' effect)	**The effect of adhering to intervention…**	failures in implementing the intervention that could have affected the outcome; non-adherence to their assigned intervention by trial participants
Experimental	atezolizumab	**Comparator**	docetaxel	**Source**	
Outcome	OS @ 1% PD-L1	**Results**	HR 0.77 95%CI 0.64-0.92	**Weight**	1

Domain	**Signalling question**	**Response**	**Comments**
Bias arising from the randomization process	1.1 Was the allocation sequence random?	Y	
	1.2 Was the allocation sequence concealed until participants were enrolled and assigned to interventions?	Y	
	1.3 Did baseline differences between intervention groups suggest a problem with the randomization process?	N	
	Risk of bias judgement	**Low**	
Bias due to deviations from intended interventions	2.1 Were participants aware of their assigned intervention during the trial?	Y	
	2.2 Were carers and people delivering the interventions aware of participants' assigned intervention during the trial?	Y	
	2.3. [If applicable:] If Y/PY/NI to 2.1 or 2.2: Were important non-protocol interventions balanced across intervention groups?	NA	
	2.4. [If applicable:] Were there failures in implementing the intervention that could have affected the outcome?	PN	30 subjects dropped prior to dosing with docetaxel versus a fraction for pembro
	2.5. [If applicable:] Was there non-adherence to the assigned intervention regimen that could have affected participants' outcomes?	N	
	2.6. If N/PN/NI to 2.3, or Y/PY/NI to 2.4 or 2.5: Was an appropriate analysis used to estimate the effect of adhering to the intervention?	NA	
	Risk of bias judgement	**Low**	
Bias due to missing outcome data	3.1 Were data for this outcome available for all, or nearly all, participants randomized?	Y	
	3.2 If N/PN/NI to 3.1: Is there evidence that result was not biased by missing outcome data?	NA	
	3.3 If N/PN to 3.2: Could missingness in the outcome depend on its true value?	NA	
	3.4 If Y/PY/NI to 3.3: Is it likely that missingness in the outcome depended on its true value?	NA	
	Risk of bias judgement	**Low**	
Bias in measurement of the outcome	4.1 Was the method of measuring the outcome inappropriate?	N	
	4.2 Could measurement or ascertainment of the outcome have differed between intervention groups?	N	
	4.3 Were outcome assessors aware of the intervention received by study participants?	PN	
	4.4 If Y/PY/NI to 4.3: Could assessment of the outcome have been influenced by knowledge of intervention received?	NA	
	4.5 If Y/PY/NI to 4.4: Is it likely that assessment of the outcome was influenced by knowledge of intervention received?	NA	
	Risk of bias judgement	**Low**	
Bias in selection of the reported result	5.1 Were the data that produced this result analysed in accordance with a pre-specified analysis plan that was finalized before unblinded outcome data were available for analysis?	Y	
	5.2 ... multiple eligible outcome measurements (e.g. scales, definitions, time points) within the outcome domain?	N	
	5.3 ... multiple eligible analyses of the data?	N	
	Risk of bias judgement	**Low**	
Overall bias	**Risk of bias judgement**	**Low**	

Unique ID	23	**Study ID**	NCT01903993	**Assessor**	KK
Ref or Label	Fehrebacher	**Aim**	adhering to intervention (the 'per-protocol' effect)	**The effect of adhering to intervention…**	failures in implementing the intervention that could have affected the outcome; non-adherence to their assigned intervention by trial participants
Experimental	atezolizumab	**Comparator**	docetaxel	**Source**	Journal article(s) with results of the trial
Outcome	OS @ 1% TPS	**Results**	HR 0.59 95% CI 0.4-0.85	**Weight**	1

Domain	Signalling question	Response	Comments
Bias arising from the randomization process	1.1 Was the allocation sequence random?	Y	
	1.2 Was the allocation sequence concealed until participants were enrolled and assigned to interventions?	Y	
	1.3 Did baseline differences between intervention groups suggest a problem with the randomization process?	PN	Baseline characteristics were balanced between groups, except for an 11% greater proportion of female patients in the docetaxel group (35% in the atezolizumab group vs 47% in the docetaxel group;
	Risk of bias judgement	**Low**	
Bias due to deviations from intended interventions	2.1 Were participants aware of their assigned intervention during the trial?	Y	
	2.2 Were carers and people delivering the interventions aware of participants' assigned intervention during the trial?	Y	
	2.3. [If applicable:] If Y/PY/NI to 2.1 or 2.2: Were important non-protocol interventions balanced across intervention groups?	NA	
	2.4. [If applicable:] Were there failures in implementing the intervention that could have affected the outcome?	N	
	2.5. [If applicable:] Was there non-adherence to the assigned intervention regimen that could have affected participants' outcomes?	N	
	2.6. If N/PN/NI to 2.3, or Y/PY/NI to 2.4 or 2.5: Was an appropriate analysis used to estimate the effect of adhering to the intervention?	NA	
	Risk of bias judgement	**Low**	
Bias due to missing outcome data	3.1 Were data for this outcome available for all, or nearly all, participants randomized?	Y	
	3.2 If N/PN/NI to 3.1: Is there evidence that result was not biased by missing outcome data?	NA	
	3.3 If N/PN to 3.2: Could missingness in the outcome depend on its true value?	NA	
	3.4 If Y/PY/NI to 3.3: Is it likely that missingness in the outcome depended on its true value?	NA	
	Risk of bias judgement	**Low**	
Bias in measurement of the outcome	4.1 Was the method of measuring the outcome inappropriate?	N	
	4.2 Could measurement or ascertainment of the outcome have differed between intervention groups?	N	
	4.3 Were outcome assessors aware of the intervention received by study participants?	PY	
	4.4 If Y/PY/NI to 4.3: Could assessment of the outcome have been influenced by knowledge of intervention received?	PN	
	4.5 If Y/PY/NI to 4.4: Is it likely that assessment of the outcome was influenced by knowledge of intervention received?	NA	
	Risk of bias judgement	**Low**	
Bias in selection of the reported result	5.1 Were the data that produced this result analysed in accordance with a pre-specified analysis plan that was finalized before unblinded outcome data were available for analysis?	Y	
	5.2 ... multiple eligible outcome measurements (e.g. scales, definitions, time points) within the outcome domain?	N	
	5.3 ... multiple eligible analyses of the data?	N	
	Risk of bias judgement	**Low**	
Overall bias	**Risk of bias judgement**	**Low**	

Appendix C

Overall Survival Results

Table C-1

Overall Survival by PD-L1 Positivity using 1% Threshold Across All Tumors

```
.
. meta summarize, subgroup(PDL1_Pos)

  Effect-size label:  HR
        Effect size:  OS_HR
          Std. Err.:  _meta_se
        Study label:  Study

Subgroup meta-analysis summary              Number of studies =     28
Random-effects model
Method: REML
Group: PDL1_Pos
```

Study	HR	[95% Conf.	Interval]	% Weight
Group: No				
Imblaze 370	1.810	0.770	2.850	0.32
JAVELIN Gastric 300	1.220	0.855	1.585	2.06
KEYNOTE-061	1.200	0.830	1.570	2.01
CheckMate 141	0.730	0.430	1.030	2.72
Checkmate 067	0.770	0.525	1.015	3.52
KEYNOTE 006	0.790	0.440	1.140	2.19
CheckMate 078	0.750	0.465	1.035	2.92
CheckMate 017	0.580	0.305	0.855	3.05
CheckMate 57	0.900	0.610	1.190	2.85
OAK	0.840	0.670	1.010	5.04
POPLAR	1.040	0.475	1.605	1.00
CheckMate 25	0.770	0.585	0.955	4.69
theta	0.829	0.749	0.909	
Group: Yes				
Imblaze 370	0.800	0.310	1.290	1.28
JAVELIN Gastric 300	0.940	0.450	1.430	1.28
KEYNOTE-061	0.820	0.635	1.005	4.69
CheckMate 141	0.550	0.355	0.745	4.47
Checkmate 067	0.530	0.380	0.680	5.53
KEYNOTE 006	0.670	0.515	0.825	5.41
JAVELIN Lung 200	0.900	0.700	1.100	4.36
CheckMate 078	0.620	0.410	0.830	4.16
CheckMate 017	0.690	0.390	0.990	2.72
CheckMate 26	1.070	0.835	1.305	3.69
CheckMate 57	0.590	0.395	0.785	4.47
KEYNOTE 010	0.660	0.560	0.760	6.83
KEYNOTE-042	0.810	0.700	0.920	6.57
OAK	0.770	0.630	0.910	5.79
POPLAR	0.590	0.365	0.815	3.87
CheckMate 25	0.790	0.470	1.110	2.49
theta	0.718	0.646	0.790	
Overall				
theta	0.759	0.699	0.820	

Heterogeneity summary

Group	df	Q	P > Q	tau2	% I2	H2
No	11	16.99	0.108	0.000	0.00	1.00
Yes	15	31.53	0.007	0.011	55.56	2.25
Overall	27	54.81	0.001	0.011	49.97	2.00

Test of group differences: Q_b = chi2(1) = 4.08 Prob > Q_b = 0.043

Table C-2

Overall Survival by PD-L1 Positivity using 5% Threshold Across All Tumors

```
. meta summarize, subgroup(PDL1_Pos)

  Effect-size label:  HR
        Effect size:  OS_HR
          Std. Err.:  _meta_se
        Study label:  Study

Subgroup meta-analysis summary            Number of studies =     16
Random-effects model
Method: REML
Group: PDL1_Pos
```

Study	HR	[95% Conf.	Interval]	% Weight
Group: No				
CheckMate 57	1.010	0.725	1.295	5.78
CheckMate 017	0.700	0.425	0.975	5.94
CheckMate 141	0.810	0.480	1.140	5.09
Checkmate 067	0.620	0.470	0.770	8.09
CheckMate 066	0.480	0.285	0.675	7.32
CheckMate 037	1.150	0.750	1.550	4.18
theta	0.758	0.567	0.949	
Group: Yes				
CheckMate 57	0.430	0.265	0.595	7.84
CheckMate 017	0.530	0.240	0.820	5.70
CheckMate 26	1.020	0.770	1.270	6.36
OAK	0.640	0.465	0.815	7.67
POPLAR	0.540	0.260	0.820	5.86
CheckMate 141	0.500	0.235	0.765	6.11
Checkmate 067	0.630	0.360	0.900	6.02
Imvigor211	0.870	0.580	1.160	5.70
CheckMate 066	0.300	0.075	0.525	6.79
CheckMate 037	0.730	0.430	1.030	5.54
theta	0.610	0.477	0.744	
Overall				
theta	0.663	0.554	0.772	

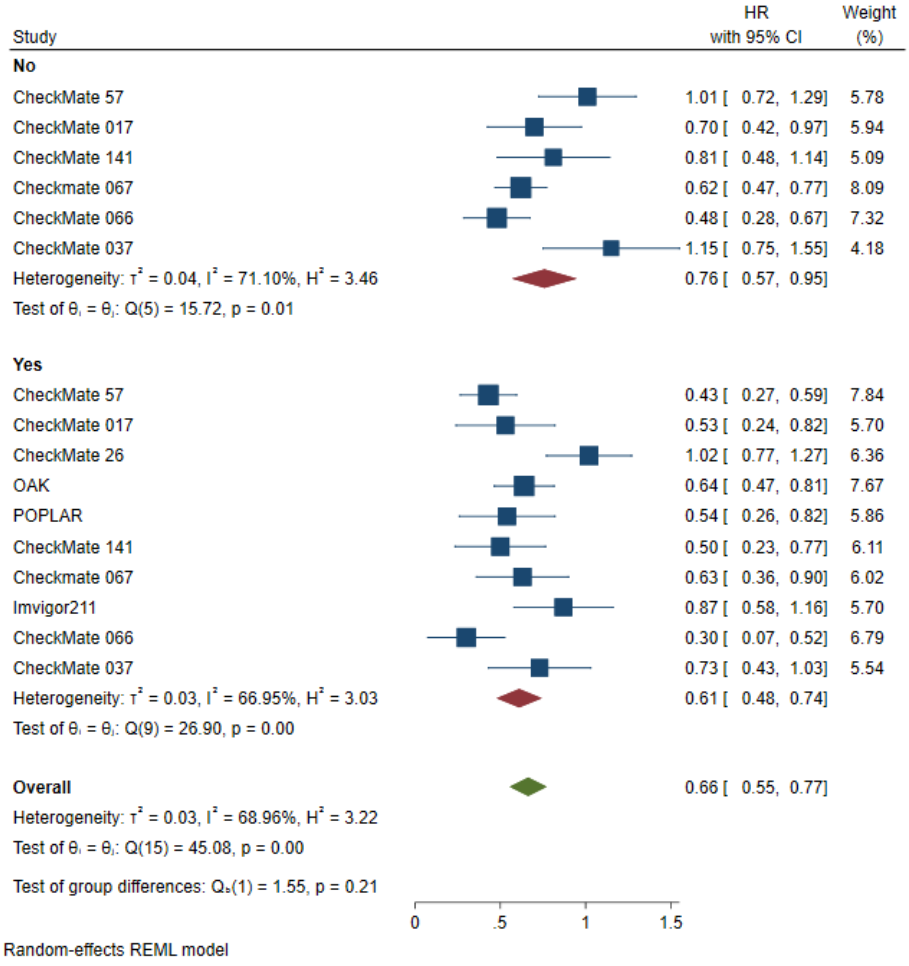
Heterogeneity summary

Group	df	Q	P > Q	tau2	% I2	H2
No	5	15.72	0.008	0.038	71.10	3.46
Yes	9	26.90	0.001	0.030	66.95	3.03
Overall	15	45.08	0.000	0.032	68.96	3.22

Test of group differences: Q_b = chi2(1) = 1.55 Prob > Q_b = 0.213

Figure C-1

Overall Survival by PD-L1 Positivity using 5% Threshold Across All Tumors



Legend. No = Negative for PD-L1 at 1% threshold Yes = Positive for PD-L1 at 5% threshold

Table C-3

Overall Survival by PD-L1 Positivity using 10% Threshold Across All Tumors

```
. meta summarize, subgroup(PDL1_Pos)

  Effect-size label:  HR
        Effect size:  OS_HR
          Std. Err.:  _meta_se
        Study label:  Study

Subgroup meta-analysis summary            Number of studies =      9
Random-effects model
Method: REML
Group: PDL1_Pos
```

Study	HR	[95% Conf.	Interval]	% Weight
Group: No				
CheckMate 57	1.000	0.725	1.275	10.31
CheckMate 017	0.700	0.435	0.965	10.74
CheckMate 141	0.730	0.450	1.010	10.10
Checkmate 074	0.610	0.465	0.755	17.20
theta	0.737	0.571	0.903	
Group: Yes				
KEYNOTE-045	0.570	0.315	0.825	11.18
CheckMate 57	0.400	0.235	0.565	15.99
CheckMate 017	0.500	0.195	0.805	9.14
CheckMate 141	0.560	0.210	0.910	7.67
Checkmate 072	0.710	0.360	1.060	7.67
theta	0.501	0.382	0.620	
Overall				
theta	0.628	0.509	0.746	

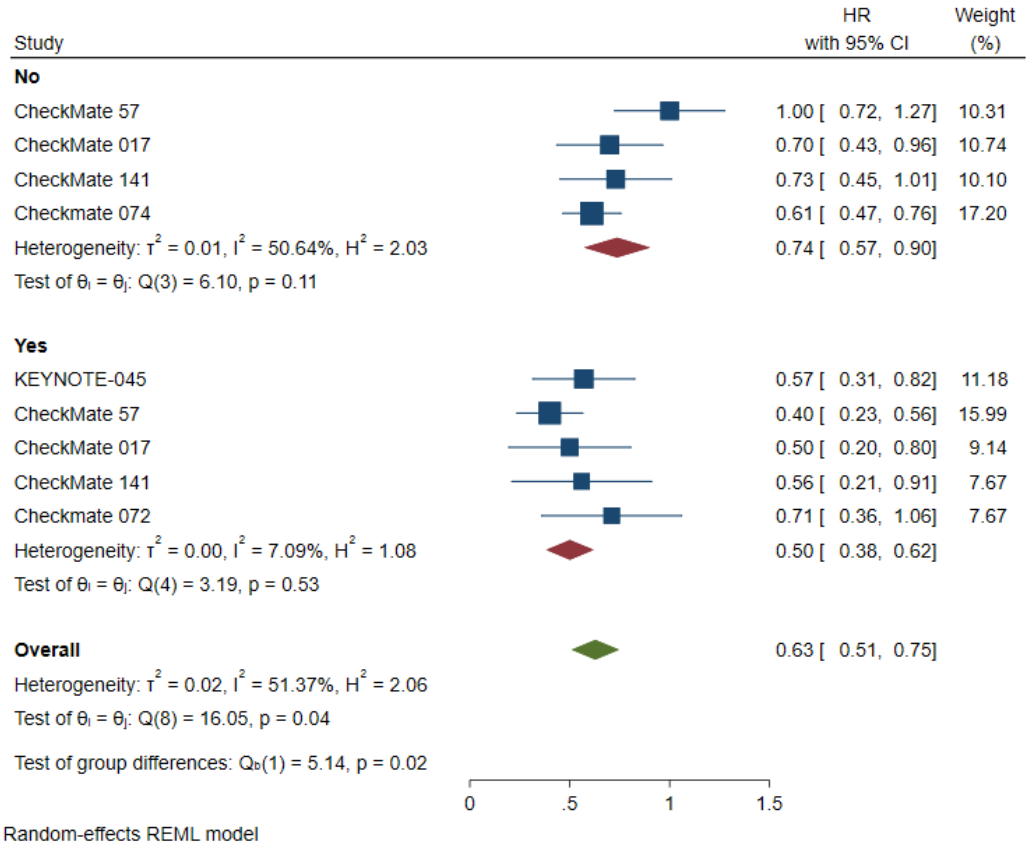
Heterogeneity summary

Group	df	Q	P > Q	tau2	% I2	H2
No	3	6.10	0.107	0.014	50.64	2.03
Yes	4	3.19	0.527	0.001	7.09	1.08
Overall	8	16.05	0.042	0.016	51.37	2.06

Test of group differences: Q_b = chi2(1) = 5.14 Prob > Q_b = 0.023

Figure C-2

Overall Survival by PD-L1 Positivity using 10% Threshold Across All Tumors



Legend. No = Negative for PD-L1 at 1% threshold Yes = Positive for PD-L1 at 10% threshold

Table C-4

Overall Survival by PD-L1 Positivity using 50% Threshold Across All Tumors

```
. meta summarize, subgroup(PDL1_Pos)

  Effect-size label:  HR
        Effect size:  OS_HR
          Std. Err.:  _meta_se
        Study label:  Study

Subgroup meta-analysis summary            Number of studies =      9
Random-effects model
Method: REML
Group: PDL1_Pos
```

Study	HR	[95% Conf.	Interval]	% Weight
Group: No				
KEYNOTE-040	0.930	0.710	1.150	10.84
theta	0.930	0.710	1.150	
Group: Yes				
JAVELIN Lung 200	0.670	0.480	0.860	12.29
CheckMate 26	0.900	0.570	1.230	6.85
KEYNOTE 010	0.500	0.375	0.625	15.83
KEYNOTE-024	0.630	0.435	0.825	12.04
KEYNOTE-040	0.530	0.300	0.760	10.39
KEYNOTE-042	0.690	0.545	0.835	14.71
OAK	0.450	0.260	0.640	12.29
POPLAR	0.490	0.065	0.915	4.77
theta	0.594	0.509	0.679	
Overall				
theta	0.635	0.529	0.741	

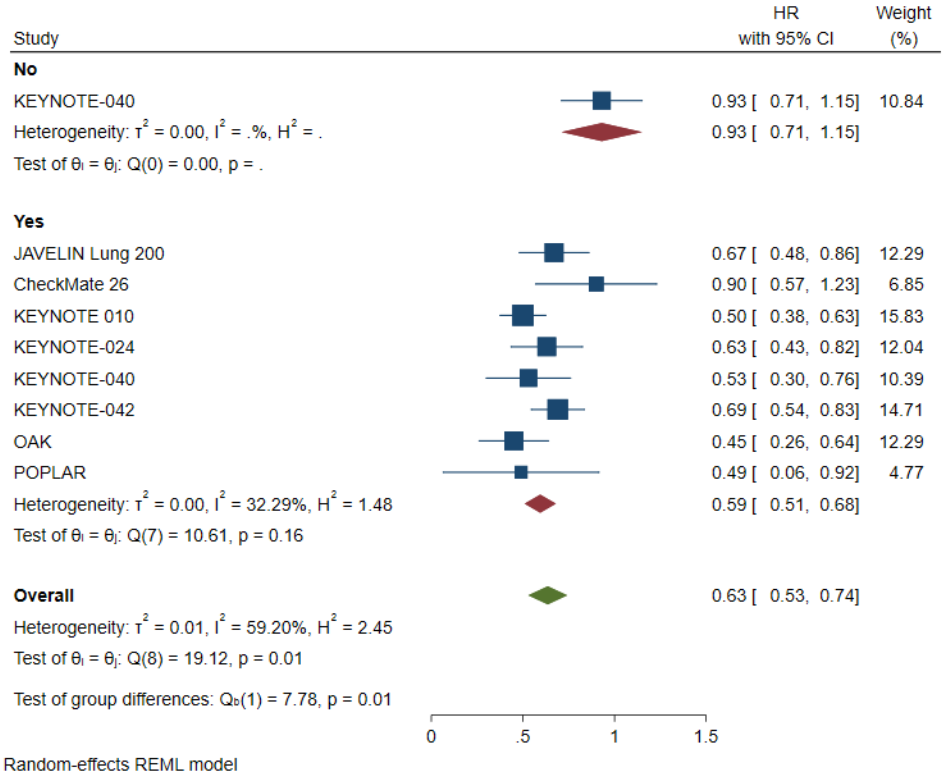
Heterogeneity summary

Group	df	Q	P > Q	tau2	% I2	H2
No	0	0.00	.	0.000	.	.
Yes	7	10.61	0.157	0.005	32.29	1.48
Overall	8	19.12	0.014	0.014	59.20	2.45

Test of group differences: Q_b = chi2(1) = 7.78 Prob > Q_b = 0.005

Figure C-3

Overall Survival by PD-L1 Positivity using 50% Threshold Across All Tumors



Legend. No = Negative for PD-L1 at 1% threshold Yes = Positive for PD-L1 at 50% threshold

Overall Survival Sub-Group Analysis

Table C-5

Overall Survival for PD-L1 Negative Patients using 1% Threshold Across All Tumors grouped by PD-1 inhibitor versus PD-L1 Inhibitor

```
. meta summarize, subgroup(PD1Inhibitor)

  Effect-size label:  HR
        Effect size:  OS_HR
          Std. Err.:  _meta_se
        Study label:  Study

Subgroup meta-analysis summary              Number of studies =     12
Random-effects model
Method: REML
Group: PD1Inhibitor
```

Study	HR	[95% Conf.	Interval]	% Weight
Group: No				
JAVELIN Gastric 300	1.220	0.855	1.585	4.81
Imblaze 370	1.810	0.770	2.850	0.59
OAK	0.840	0.670	1.010	22.17
POPLAR	1.040	0.475	1.605	2.01
theta	1.054	0.762	1.345	
Group: Yes				
CheckMate 078	0.750	0.465	1.035	7.89
CheckMate 141	0.730	0.430	1.030	7.12
CheckMate 017	0.580	0.305	0.855	8.47
CheckMate 25	0.770	0.585	0.955	18.72
CheckMate 57	0.900	0.610	1.190	7.62
Checkmate 067	0.770	0.525	1.015	10.68
KEYNOTE 006	0.790	0.440	1.140	5.23
KEYNOTE-061	1.200	0.830	1.570	4.68
theta	0.785	0.690	0.880	
Overall				
theta	0.829	0.749	0.909	

Heterogeneity summary

Group	df	Q	P > Q	tau2	% I2	H2
No	3	6.39	0.094	0.041	49.38	1.98
Yes	7	7.80	0.351	0.000	0.00	1.00
Overall	11	16.99	0.108	0.000	0.00	1.00

Test of group differences: Q_b = chi2(1) = 2.95 Prob > Q_b = 0.086

Table C-6

Overall Survival for PD-L1 Negative Patients using 1% Threshold Across All Tumors grouped by Follow-up Duration (<18mon versus ≥18mon)

```
. meta summarize, subgroup(FUDuration18months)

  Effect-size label:  HR
        Effect size:  OS_HR
          Std. Err.:  _meta_se
        Study label:  Study

Subgroup meta-analysis summary              Number of studies =     12
Random-effects model
Method: REML
Group: FUDuration18months
```

Study	HR	[95% Conf.	Interval]	% Weight
Group: No				
CheckMate 078	0.750	0.465	1.035	7.89
CheckMate 017	0.580	0.305	0.855	8.47
CheckMate 57	0.900	0.610	1.190	7.62
Imblaze 370	1.810	0.770	2.850	0.59
KEYNOTE-061	1.200	0.830	1.570	4.68
POPLAR	1.040	0.475	1.605	2.01
theta	0.902	0.673	1.132	
Group: Yes				
JAVELIN Gastric 300	1.220	0.855	1.585	4.81
CheckMate 141	0.730	0.430	1.030	7.12
CheckMate 25	0.770	0.585	0.955	18.72
Checkmate 067	0.770	0.525	1.015	10.68
KEYNOTE 006	0.790	0.440	1.140	5.23
OAK	0.840	0.670	1.010	22.17
theta	0.821	0.725	0.918	
Overall				
theta	0.829	0.749	0.909	

Heterogeneity summary

Group	df	Q	P > Q	tau2	% I2	H2
No	5	11.43	0.043	0.041	54.35	2.19
Yes	5	5.48	0.360	0.000	0.00	1.00
Overall	11	16.99	0.108	0.000	0.00	1.00

Test of group differences: Q_b = chi2(1) = 0.40 Prob > Q_b = 0.525

Table C-**7**

Overall Survival for PD-L1 Negative Patients using 1% Threshold Across All Tumors grouped by Line of Therapy (1 Line versus Later Line)

```
. meta summarize, subgroup(FirstLine)

  Effect-size label:  HR
        Effect size:  OS_HR
          Std. Err.:  _meta_se
        Study label:  Study

Subgroup meta-analysis summary              Number of studies =     12
Random-effects model
Method: REML
Group: FirstLine
```

Study	HR	[95% Conf.	Interval]	% Weight
Group: No				
CheckMate 078	0.750	0.465	1.035	7.89
JAVELIN Gastric 300	1.220	0.855	1.585	4.81
CheckMate 141	0.730	0.430	1.030	7.12
CheckMate 017	0.580	0.305	0.855	8.47
CheckMate 25	0.770	0.585	0.955	18.72
CheckMate 57	0.900	0.610	1.190	7.62
Imblaze 370	1.810	0.770	2.850	0.59
KEYNOTE-061	1.200	0.830	1.570	4.68
OAK	0.840	0.670	1.010	22.17
POPLAR	1.040	0.475	1.605	2.01
theta	0.863	0.740	0.985	
Group: Yes				
Checkmate 067	0.770	0.525	1.015	10.68
KEYNOTE 006	0.790	0.440	1.140	5.23
theta	0.777	0.576	0.977	
Overall				
theta	0.829	0.749	0.909	

Heterogeneity summary

Group	df	Q	P > Q	tau2	% I2	H2
No	9	16.67	0.054	0.014	39.72	1.66
Yes	1	0.01	0.927	0.000	0.00	1.00
Overall	11	16.99	0.108	0.000	0.00	1.00

```
Test of group differences: Q_b = chi2(1) = 0.51          Prob > Q_b = 0.473
```

Table C-8

Overall Survival for PD-L1 Positive Patients using 1% Threshold Across All Tumors grouped by PD-1 inhibitor versus PD-L1 Inhibitor

```
. meta summarize, subgroup(PD1Inhibitor)

  Effect-size label:  HR
        Effect size:  OS_HR
          Std. Err.:  _meta_se
        Study label:  Study

Subgroup meta-analysis summary              Number of studies =     16
Random-effects model
Method: REML
Group: PD1Inhibitor
```

Study	HR	[95% Conf.	Interval]	% Weight
Group: No				
JAVELIN Lung 200	0.900	0.700	1.100	6.43
JAVELIN Gastric 300	0.940	0.450	1.430	1.85
Imblaze 370	0.800	0.310	1.290	1.85
OAK	0.770	0.630	0.910	8.62
POPLAR	0.590	0.365	0.815	5.68
theta	0.776	0.653	0.899	
Group: Yes				
CheckMate 078	0.620	0.410	0.830	6.12
CheckMate 141	0.550	0.355	0.745	6.59
CheckMate 017	0.690	0.390	0.990	3.97
CheckMate 25	0.790	0.470	1.110	3.62
CheckMate 26	1.070	0.835	1.305	5.41
CheckMate 57	0.590	0.395	0.785	6.59
Checkmate 067	0.530	0.380	0.680	8.22
KEYNOTE 006	0.670	0.515	0.825	8.02
KEYNOTE 010	0.660	0.560	0.760	10.25
KEYNOTE-061	0.820	0.635	1.005	6.93
KEYNOTE-042	0.810	0.700	0.920	9.84
theta	0.700	0.613	0.786	
Overall				
theta	0.718	0.646	0.790	

Heterogeneity summary

Group	df	Q	P > Q	tau2	% I2	H2
No	4	4.55	0.337	0.005	23.92	1.31
Yes	10	25.02	0.005	0.012	62.43	2.66
Overall	15	31.53	0.007	0.011	55.56	2.25

Test of group differences: Q_b = chi2(1) = 0.98 Prob > Q_b = 0.322

Table C-9

Overall Survival for PD-L1 Positive Patients using 1% Threshold Across All Tumors grouped by Follow-up Duration (<18mon versus ≥18mon)

```
. meta summarize, subgroup(FU18mon)

  Effect-size label:  HR
        Effect size:  OS_HR
          Std. Err.:  _meta_se
        Study label:  Study

Subgroup meta-analysis summary              Number of studies =     16
Random-effects model
Method: REML
Group: FU18mon
```

Study	HR	[95% Conf.	Interval]	% Weight
Group: No				
CheckMate 078	0.620	0.410	0.830	6.12
CheckMate 017	0.690	0.390	0.990	3.97
CheckMate 26	1.070	0.835	1.305	5.41
CheckMate 57	0.590	0.395	0.785	6.59
Imblaze 370	0.800	0.310	1.290	1.85
KEYNOTE 010	0.660	0.560	0.760	10.25
KEYNOTE-061	0.820	0.635	1.005	6.93
KEYNOTE-042	0.810	0.700	0.920	9.84
POPLAR	0.590	0.365	0.815	5.68
theta	0.732	0.634	0.830	
Group: Yes				
JAVELIN Lung 200	0.900	0.700	1.100	6.43
JAVELIN Gastric 300	0.940	0.450	1.430	1.85
CheckMate 141	0.550	0.355	0.745	6.59
CheckMate 25	0.790	0.470	1.110	3.62
Checkmate 067	0.530	0.380	0.680	8.22
KEYNOTE 006	0.670	0.515	0.825	8.02
OAK	0.770	0.630	0.910	8.62
theta	0.702	0.588	0.816	
Overall				
theta	0.718	0.646	0.790	

Heterogeneity summary

Group	df	Q	P > Q	tau2	% I2	H2
No	8	17.53	0.025	0.011	57.79	2.37
Yes	6	13.28	0.039	0.012	56.24	2.29
Overall	15	31.53	0.007	0.011	55.56	2.25

Test of group differences: Q_b = chi2(1) = 0.15 Prob > Q_b = 0.695

Table C-10

Overall Survival for PD-L1 Positive Patients using 1% Threshold Across All Tumors grouped by Line of Therapy (1 Line versus Later Line)

```
. meta summarize, subgroup(FirstLine)

  Effect-size label:  HR
        Effect size:  OS_HR
          Std. Err.:  _meta_se
        Study label:  Study

Subgroup meta-analysis summary              Number of studies =     16
Random-effects model
Method: REML
Group: FirstLine
```

Study	HR	[95% Conf.	Interval]	% Weight
Group: No				
JAVELIN Lung 200	0.900	0.700	1.100	6.43
CheckMate 078	0.620	0.410	0.830	6.12
JAVELIN Gastric 300	0.940	0.450	1.430	1.85
CheckMate 141	0.550	0.355	0.745	6.59
CheckMate 017	0.690	0.390	0.990	3.97
CheckMate 25	0.790	0.470	1.110	3.62
CheckMate 57	0.590	0.395	0.785	6.59
Imblaze 370	0.800	0.310	1.290	1.85
KEYNOTE 010	0.660	0.560	0.760	10.25
KEYNOTE-061	0.820	0.635	1.005	6.93
OAK	0.770	0.630	0.910	8.62
POPLAR	0.590	0.365	0.815	5.68
theta	0.702	0.634	0.770	
Group: Yes				
CheckMate 26	1.070	0.835	1.305	5.41
Checkmate 067	0.530	0.380	0.680	8.22
KEYNOTE 006	0.670	0.515	0.825	8.02
KEYNOTE-042	0.810	0.700	0.920	9.84
theta	0.759	0.548	0.970	
Overall				
theta	0.718	0.646	0.790	

Heterogeneity summary

Group	df	Q	P > Q	tau2	% I2	H2
No	11	13.39	0.269	0.003	23.73	1.31
Yes	3	17.44	0.001	0.039	86.61	7.47
Overall	15	31.53	0.007	0.011	55.56	2.25

Test of group differences: Q_b = chi2(1) = 0.25 Prob > Q_b = 0.618

Sensitivity Analysis

Table C-11

Overall Survival by PD-L1 Positivity using 1% Threshold Across All Tumors with removal of CHECKMATE-026 study

```
.
. meta summarize, subgroup(PDL1_Pos)

  Effect-size label:  HR
        Effect size:  OS_HR
          Std. Err.:  _meta_se
        Study label:  Study

Subgroup meta-analysis summary              Number of studies =     27
Random-effects model
Method: REML
Group: PDL1_Pos
```

Study	HR	[95% Conf.	Interval]	% Weight
Group: No				
CheckMate 078	0.750	0.465	1.035	2.82
JAVELIN Gastric 300	1.220	0.855	1.585	1.92
CheckMate 141	0.730	0.430	1.030	2.61
CheckMate 017	0.580	0.305	0.855	2.97
CheckMate 25	0.770	0.585	0.955	4.92
CheckMate 57	0.900	0.610	1.190	2.75
Checkmate 067	0.770	0.525	1.015	3.49
Imblaze 370	1.810	0.770	2.850	0.28
KEYNOTE 006	0.790	0.440	1.140	2.05
KEYNOTE-061	1.200	0.830	1.570	1.87
OAK	0.840	0.670	1.010	5.37
POPLAR	1.040	0.475	1.605	0.89
theta	0.829	0.749	0.909	
Group: Yes				
JAVELIN Lung 200	0.900	0.700	1.100	4.51
CheckMate 078	0.620	0.410	0.830	4.26
JAVELIN Gastric 300	0.940	0.450	1.430	1.16
CheckMate 141	0.550	0.355	0.745	4.64
CheckMate 017	0.690	0.390	0.990	2.61
CheckMate 25	0.790	0.470	1.110	2.37
CheckMate 57	0.590	0.395	0.785	4.64
Checkmate 067	0.530	0.380	0.680	6.04
Imblaze 370	0.800	0.310	1.290	1.16
KEYNOTE 006	0.670	0.515	0.825	5.87
KEYNOTE 010	0.660	0.560	0.760	7.98
KEYNOTE-061	0.820	0.635	1.005	4.92
KEYNOTE-042	0.810	0.700	0.920	7.57
OAK	0.770	0.630	0.910	6.40
POPLAR	0.590	0.365	0.815	3.91
theta	0.698	0.634	0.761	
Overall				
theta	0.743	0.688	0.799	

Heterogeneity summary

Group	df	Q	P > Q	tau2	% I2	H2
No	11	16.99	0.108	0.000	0.00	1.00
Yes	14	22.31	0.072	0.006	41.07	1.70
Overall	26	47.00	0.007	0.008	40.36	1.68

Test of group differences: Q_b = chi2(1) = 6.39 Prob > Q_b = 0.011

Table C-12

Overall Survival by PD-L1 Positivity using 5% Threshold Across All Tumors with removal of CHECKMATE-026 study

```
. meta summarize, subgroup(PDL1_Pos)

  Effect-size label:  HR
        Effect size:  OS_HR
          Std. Err.:  _meta_se
        Study label:  Study

Subgroup meta-analysis summary            Number of studies =     15
Random-effects model
Method: REML
Group: PDL1_Pos
```

Study	HR	[95% Conf.	Interval]	% Weight
Group: No				
CheckMate 141	0.810	0.480	1.140	5.20
CheckMate 017	0.700	0.425	0.975	6.24
CheckMate 037	1.150	0.750	1.550	4.14
CheckMate 57	1.010	0.725	1.295	6.03
CheckMate 066	0.480	0.285	0.675	8.05
Checkmate 067	0.620	0.470	0.770	9.15
theta	0.758	0.567	0.949	
Group: Yes				
CheckMate 141	0.500	0.235	0.765	6.45
CheckMate 017	0.530	0.240	0.820	5.94
CheckMate 037	0.730	0.430	1.030	5.74
CheckMate 57	0.430	0.265	0.595	8.78
CheckMate 066	0.300	0.075	0.525	7.33
Checkmate 067	0.630	0.360	0.900	6.34
Imvigor211	0.870	0.580	1.160	5.94
OAK	0.640	0.465	0.815	8.54
POPLAR	0.540	0.260	0.820	6.14
theta	0.559	0.450	0.667	
Overall				
theta	0.634	0.532	0.737	

Heterogeneity summary

Group	df	Q	P > Q	tau2	% I2	H2
No	5	15.72	0.008	0.038	71.10	3.46
Yes	8	14.34	0.073	0.012	44.63	1.81
Overall	14	34.95	0.001	0.024	62.16	2.64

Test of group differences: Q_b = chi2(1) = 3.17 Prob > Q_b = 0.075

Table C-13

Overall Survival by PD-L1 Positivity using 1% Threshold in NSCLC patients

```
. meta summarize, subgroup(PDL1_Pos)

  Effect-size label:  HR
        Effect size:  OS_HR
          Std. Err.:  _meta_se
        Study label:  Study

Subgroup meta-analysis summary            Number of studies =     14
Random-effects model
Method: REML
Group: PDL1_Pos
```

Study	HR	[95% Conf.	Interval]	% Weight
Group: No				
CheckMate 078	0.750	0.465	1.035	4.75
CheckMate 017	0.580	0.305	0.855	5.00
CheckMate 57	0.900	0.610	1.190	4.63
OAK	0.840	0.670	1.010	8.88
POPLAR	1.040	0.475	1.605	1.52
theta	0.797	0.682	0.913	
Group: Yes				
JAVELIN Lung 200	0.900	0.700	1.100	7.50
CheckMate 078	0.620	0.410	0.830	7.09
CheckMate 017	0.690	0.390	0.990	4.41
CheckMate 26	1.070	0.835	1.305	6.18
CheckMate 57	0.590	0.395	0.785	7.71
KEYNOTE 010	0.660	0.560	0.760	12.96
KEYNOTE-042	0.810	0.700	0.920	12.33
OAK	0.770	0.630	0.910	10.51
POPLAR	0.590	0.365	0.815	6.53
theta	0.743	0.650	0.836	
Overall				
theta	0.755	0.682	0.828	

Heterogeneity summary

Group	df	Q	P > Q	tau2	% I2	H2
No	4	3.94	0.415	0.000	0.00	1.00
Yes	8	19.56	0.012	0.011	61.87	2.62
Overall	13	24.35	0.028	0.008	47.49	1.90

```
Test of group differences: Q_b = chi2(1) = 0.52           Prob > Q_b = 0.470

.
```

Progression Free Survival Results

Table C-14

Progression Free Survival by PD-L1 Positivity using 1% Threshold Across All Tumors

```
. meta summarize, subgroup(PDL1_Pos)

  Effect-size label:  HR
        Effect size:  PFS_HR
          Std. Err.:  _meta_se
        Study label:  Study

Subgroup meta-analysis summary              Number of studies =     26
Random-effects model
Method: REML
Group: PDL1_Pos
```

Study	HR	[95% Conf.	Interval]	% Weight
Group: No				
JAVELIN Gastric 300	1.950	1.355	2.545	2.53
CheckMate 141	1.130	0.650	1.610	3.06
CheckMate 017	0.660	0.375	0.945	4.09
CheckMate 57	1.190	0.825	1.555	3.66
Checkmate 067	0.590	0.410	0.770	4.59
Imblaze 370	2.170	1.030	3.310	1.09
KEYNOTE-002	0.600	0.320	0.880	4.11
KEYNOTE 006	0.870	0.510	1.230	3.68
KEYNOTE-061	2.050	1.405	2.695	2.33
OAK	1.110	0.905	1.315	4.48
POPLAR	1.120	0.595	1.645	2.84
theta	1.120	0.824	1.416	
Group: Yes				
JAVELIN Lung 200	1.010	0.770	1.250	4.32
JAVELIN Gastric 300	1.750	0.755	2.745	1.34
CheckMate 141	0.590	0.375	0.805	4.44
CheckMate 017	0.670	0.385	0.955	4.09
CheckMate 26	1.170	0.930	1.410	4.32
CheckMate 57	0.700	0.495	0.905	4.48
Checkmate 067	0.450	0.335	0.565	4.83
Imblaze 370	0.710	0.315	1.105	3.49
KEYNOTE-002	0.520	0.375	0.665	4.73
KEYNOTE 006	0.540	0.435	0.645	4.86
KEYNOTE 010	0.830	0.710	0.950	4.81
KEYNOTE-061	1.270	1.000	1.540	4.16
KEYNOTE-042	1.070	0.935	1.205	4.77
OAK	0.870	0.725	1.015	4.73
POPLAR	0.850	0.585	1.115	4.19
theta	0.810	0.675	0.946	
Overall				
theta	0.911	0.777	1.046	

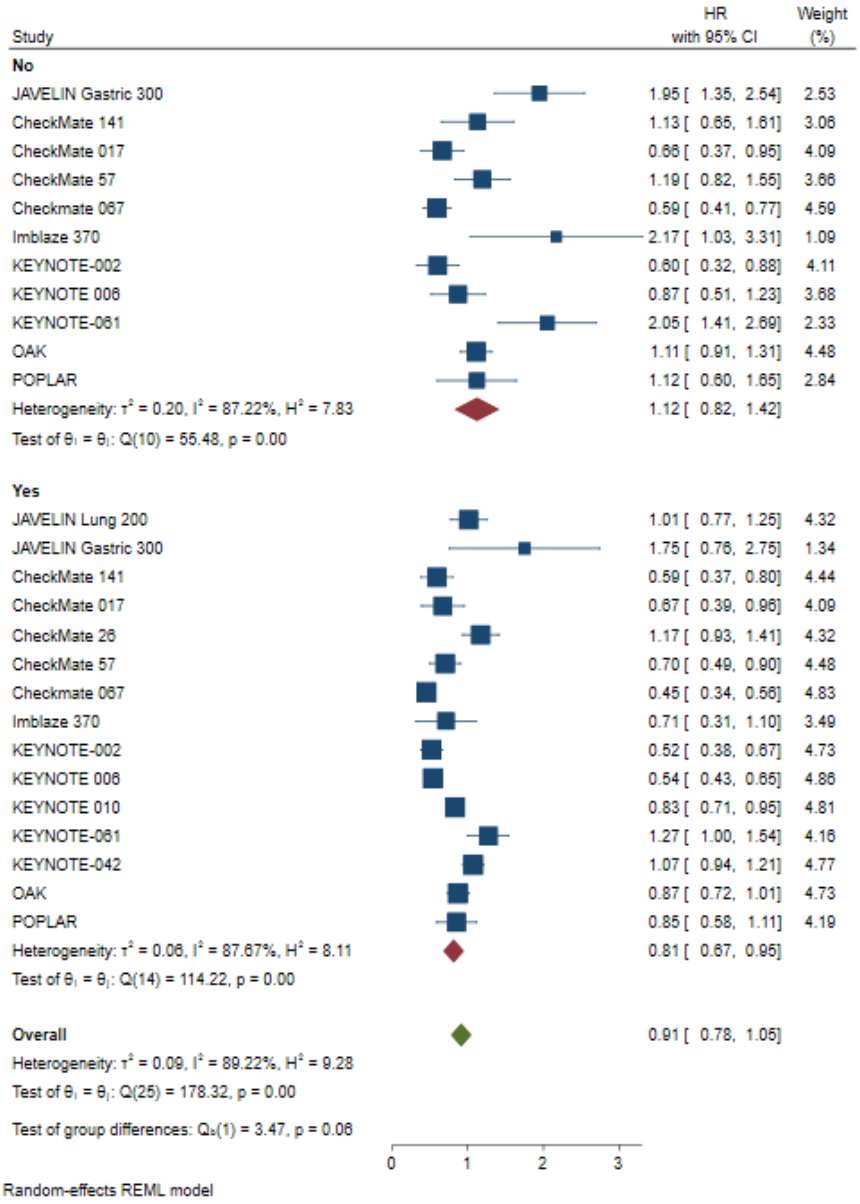
Heterogeneity summary

Group	df	Q	P > Q	tau2	% I2	H2
No	10	55.48	0.000	0.196	87.22	7.83
Yes	14	114.22	0.000	0.057	87.67	8.11
Overall	25	178.32	0.000	0.094	89.22	9.28

Test of group differences: Q_b = chi2(1) = 3.47 Prob > Q_b = 0.062

Figure C-4

PFS by PD-L1 Positivity using 1% Threshold Across All Tumors



Legend. No = Negative for PD-L1 at 1% threshold Yes = Positive for PD-L1 at 1% threshold

Table C-15

PFS by PD-L1 Positivity using 5% Threshold Across All Tumors

```
. meta summarize, subgroup(PDL1_Pos)

  Effect-size label:  HR
        Effect size:  PFS_HR
          Std. Err.:  _meta_se
        Study label:  Study

Subgroup meta-analysis summary            Number of studies =     10
Random-effects model
Method: REML
Group: PDL1_Pos
```

Study	HR	[95% Conf.	Interval]	% Weight
Group: No				
CheckMate 017	0.750	0.470	1.030	9.71
CheckMate 57	1.310	0.960	1.660	8.57
Checkmate 067	0.540	0.415	0.665	11.99
theta	0.844	0.403	1.285	
Group: Yes				
CheckMate 017	0.540	0.250	0.830	9.55
CheckMate 26	1.150	0.880	1.420	9.88
CheckMate 57	0.540	0.355	0.725	11.21
Checkmate 067	0.420	0.250	0.590	11.42
Imvigor211	1.010	0.715	1.305	9.46
OAK	0.730	0.555	0.905	11.35
POPLAR	0.720	0.255	1.185	6.85
theta	0.718	0.515	0.921	
Overall				
theta	0.751	0.572	0.930	

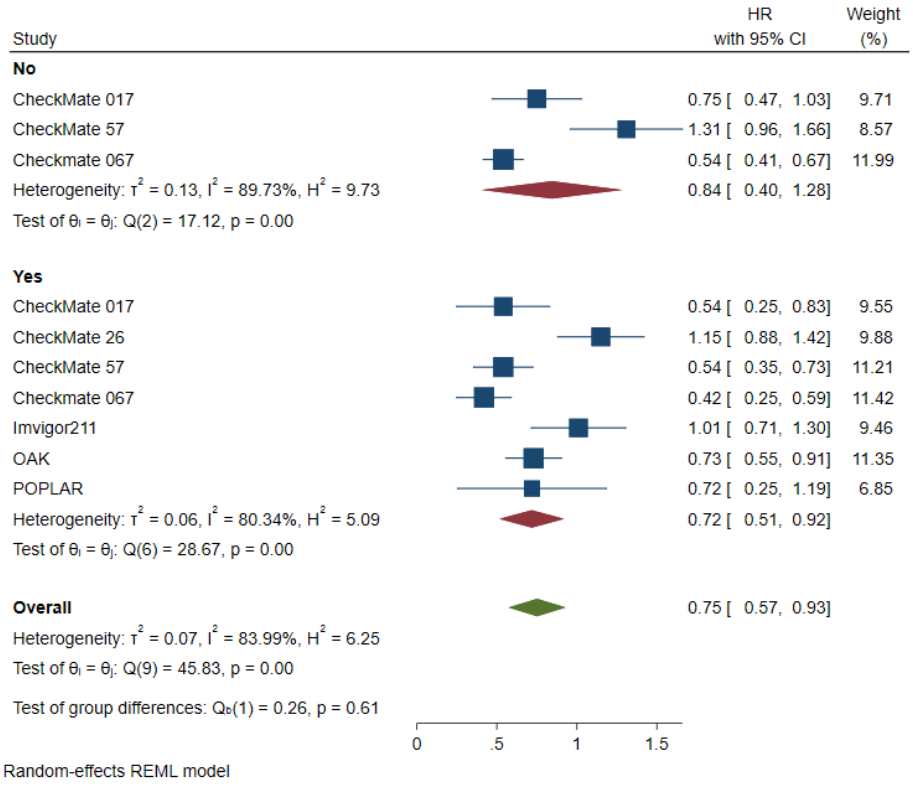
Heterogeneity summary

Group	df	Q	P > Q	tau2	% I2	H2
No	2	17.12	0.000	0.134	89.73	9.73
Yes	6	28.67	0.000	0.057	80.34	5.09
Overall	9	45.83	0.000	0.066	83.99	6.25

```
Test of group differences: Q_b = chi2(1) = 0.26           Prob > Q_b = 0.610
```

Figure C-5

PFS by PD-L1 Positivity using 5% Threshold Across All Tumors



Legend. No = Negative for PD-L1 at 1% threshold Yes = Positive for PD-L1 at 5% threshold

Table C-16

PFS by PD-L1 Positivity using 10% Threshold Across All Tumors

```
. meta summarize, subgroup(PDL1_Pos)

  Effect-size label:  HR
        Effect size:  PFS_HR
          Std. Err.:  _meta_se
        Study label:  Study

Subgroup meta-analysis summary            Number of studies =      7
Random-effects model
Method: REML
Group: PDL1_Pos
```

Study	HR	[95% Conf.	Interval]	% Weight
Group: No				
CheckMate 017	0.700	0.450	0.950	14.43
CheckMate 57	1.240	0.915	1.565	12.34
Checkmate 074	0.520	0.410	0.630	17.99
theta	0.799	0.386	1.211	
Group: Yes				
KEYNOTE-045	0.890	0.555	1.225	12.07
CheckMate 017	0.580	0.235	0.925	11.81
CheckMate 57	0.520	0.330	0.710	16.10
Checkmate 072	0.460	0.240	0.680	15.27
theta	0.573	0.420	0.727	
Overall				
theta	0.677	0.486	0.869	

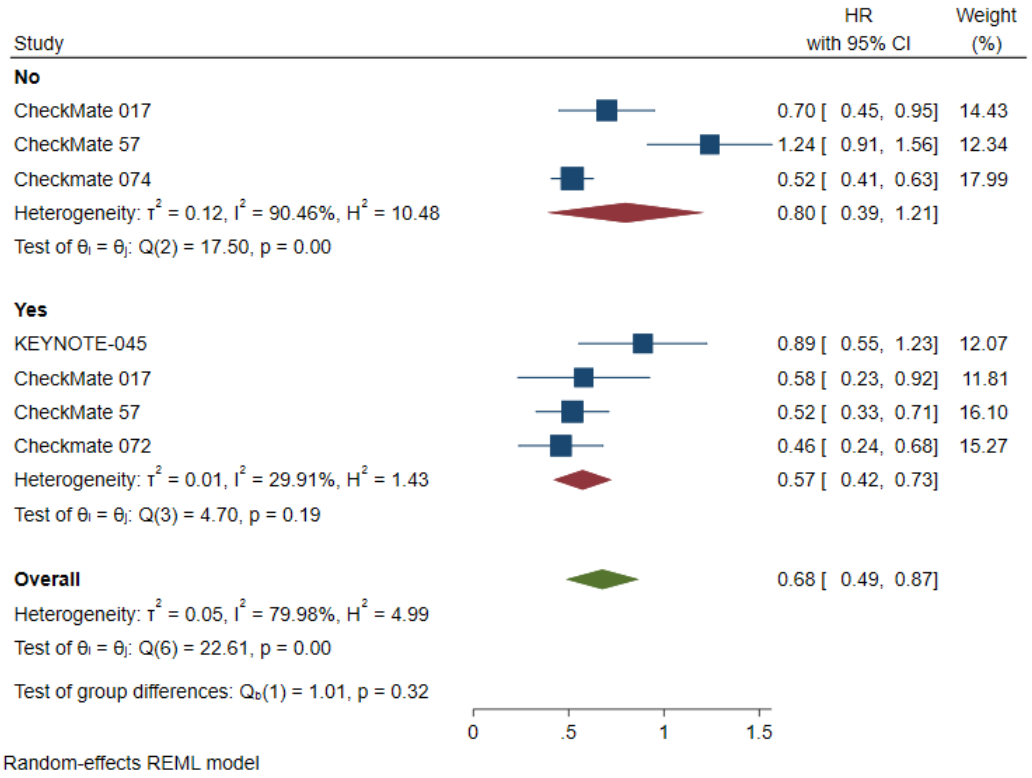
Heterogeneity summary

Group	df	Q	P > Q	tau2	% I2	H2
No	2	17.50	0.000	0.118	90.46	10.48
Yes	3	4.70	0.195	0.007	29.91	1.43
Overall	6	22.61	0.001	0.050	79.98	4.99

Test of group differences: Q_b = chi2(1) = 1.01 Prob > Q_b = 0.316

Figure C-6

PFS by PD-L1 Positivity using 10% Threshold Across All Tumors



Legend. No = Negative for PD-L1 at 1% threshold Yes = Positive for PD-L1 at 10% threshold

Table C-17

Publication Bias in Overall Survival Results for all tumors at 1% PD-L1 Threshold in Negative Patients

```
. meta bias, begg

  Effect-size label:  HR
        Effect size:  OS_HR
          Std. Err.:  _meta_se
(1 missing value generated)
(1 missing value generated)

Begg's test for small-study effects

Kendall's score =     30.00
    SE of score =    14.583
              z =      1.99
     Prob > |z| =    0.0467
```

Table C-18

Publication Bias in Overall Survival Results for all tumors at 1% PD-L1 Threshold in Positive Patients

```
. meta bias, begg

  Effect-size label:  HR
        Effect size:  OS_HR
          Std. Err.:  _meta_se
(1 missing value generated)
(1 missing value generated)

Begg's test for small-study effects

Kendall's score =     20.00
    SE of score =    22.166
              z =      0.86
     Prob > |z| =    0.3914
```

Table C-19

Publication Bias in PFS Results for all tumors at 1% PD-L1 Threshold in Negative Patients

```
. meta bias, begg

  Effect-size label:  Effect Size
        Effect size:  PFS_HR
          Std. Err.:  _meta_se
(2 missing values generated)
(2 missing values generated)

Begg's test for small-study effects

Kendall's score =     31.00
    SE of score =    12.845
              z =      2.34
     Prob > |z| =    0.0195
```

Table C-20

Publication Bias in PFS Results for all tumors at 1% PD-L1 Threshold in Positive Patients

```
. meta bias, begg

  Effect-size label:  HR
        Effect size:  PFS_HR
          Std. Err.:  _meta_se
(2 missing values generated)
(2 missing values generated)

Begg's test for small-study effects

Kendall's score =     27.00
    SE of score =    20.158
              z =      1.29
     Prob > |z| =    0.1971
```

References

Bang, Y. J., Ruiz, E. Y., Van Cutsem, E., Lee, K. W., Wyrwicz, L., Schenker, M., . . . et al. (2018). Phase III, randomised trial of avelumab versus physician's choice of chemotherapy as third-line treatment of patients with advanced gastric or gastro-oesophageal junction cancer: primary analysis of JAVELIN Gastric 300. *Annals of oncology : official journal of the European Society for Medical Oncology, 29*(10), 2052â□□2060. doi:10.1093/annonc/mdy264

Barlesi, F., Vansteenkiste, J., Spigel, D., Ishii, H., Garassino, M., de Marinis, F., . . . Park, K. (2018). Avelumab versus docetaxel in patients with platinum-treated advanced non-small-cell lung cancer (JAVELIN Lung 200): an open-label, randomised, phase 3 study. *Lancet Oncol, 19*(11), 1468-1479. doi:10.1016/s1470-2045(18)30673-9

Bellmunt, J., de Wit, R., Vaughn, D. J., Fradet, Y., Lee, J. L., Fong, L., . . . et al. (2017). Pembrolizumab as Second-Line Therapy for Advanced Urothelial Carcinoma. *New England Journal of Medicine, 376*(11), 1015â□□1026. doi:10.1056/NEJMoa1613683

Borghaei, H., Paz-Ares, L., Horn, L., Spigel, D. R., Steins, M., Ready, N. E., . . . Brahmer, J. R. (2015). Nivolumab versus Docetaxel in Advanced Nonsquamous Non-Small-Cell Lung Cancer. *N Engl J Med, 373*(17), 1627-1639. doi:10.1056/NEJMoa1507643

Borghaei, H., Paz-Ares, L., Horn, L., Spigel, D. R., Steins, M., Ready, N. E., . . . et al. (2015). Nivolumab versus Docetaxel in Advanced Nonsquamous Non-Small-Cell Lung Cancer. *New England Journal of Medicine, 373*(17), 1627â□□1639. doi:10.1056/NEJMoa1507643

Brahmer, J., Reckamp, K. L., Baas, P., CrinÃ², L., Eberhardt, W. E., Poddubskaya, E., . . . et al. (2015). Nivolumab versus Docetaxel in Advanced Squamous-Cell Non-Small-Cell Lung Cancer. *New England Journal of Medicine, 373*(2), 123â□□135. doi:10.1056/NEJMoa1504627

Brahmer, J., Reckamp, K. L., Baas, P., Crinò, L., Eberhardt, W. E., Poddubskaya, E., . . . Spigel, D. R. (2015). Nivolumab versus Docetaxel in Advanced Squamous-Cell Non-Small-Cell Lung Cancer. *N Engl J Med, 373*(2), 123-135. doi:10.1056/NEJMoa1504627

Cancer.net. (2019). Understanding Immunotherapy. Retrieved from https://www.cancer.net/navigating-cancer-care/how-cancer-treated/immunotherapy-and-vaccines/understanding-immunotherapy

Carbone, D. P., Reck, M., Paz-Ares, L., Creelan, B., Horn, L., Steins, M., . . . Socinski, M. A. (2017). First-line nivolumab in stage IV or recurrent non-small-cell lung cancer. *New England Journal of Medicine, 376*(25), 2415-2426. doi:10.1056/NEJMoa1613493

Carbone, D. P., Reck, M., Paz-Ares, L., Creelan, B., Horn, L., Steins, M., . . . Investigators, C. (2017). First-Line Nivolumab in Stage IV or Recurrent Non-Small-Cell Lung Cancer. *N Engl J Med, 376*(25), 2415-2426. doi:10.1056/NEJMoa1613493

CDC. (2016). Leading Causes of Death. Retrieved from http://www.cdc.gov/nchs/fastats/leading-causes-of-death.htm

Chen, D. S., Irving, B. A., & Hodi, F. S. (2012). Molecular pathways: next-generation immunotherapy--inhibiting programmed death-ligand 1 and programmed death-1. *Clin Cancer Res, 18*(24), 6580-6587. doi:10.1158/1078-0432.CCR-12-1362

Cohen, E. E. W., SouliÃ¨res, D., Le Tourneau, C., Dinis, J., Licitra, L., Ahn, M. J., . . . et al. (2019). Pembrolizumab versus methotrexate, docetaxel, or cetuximab for recurrent or metastatic head-and-neck squamous cell carcinoma (KEYNOTE-040): a randomised, open-label, phase 3 study. *Lancet (london, england), 393*(10167), 156â□□167. doi:10.1016/S0140-6736(18)31999-8

Conroy, J. M., Pabla, S., Nesline, M. K., Glenn, S. T., Papanicolau-Sengos, A., Burgher, B., . . . Morrison, C. (2019). Next generation sequencing of PD-L1 for predicting response to immune checkpoint inhibitors. *J Immunother Cancer, 7*(1), 18. doi:10.1186/s40425-018-0489-5

Davis, A. A., & Patel, V. G. (2019). The role of PD-L1 expression as a predictive biomarker: an analysis of all US Food and Drug Administration (FDA) approvals of immune checkpoint inhibitors. *J Immunother Cancer, 7*(1), 278. doi:10.1186/s40425-019-0768-9

De Sousa Linhares, A., Battin, C., Jutz, S., Leitner, J., Hafner, C., Tobias, J., . . . Steinberger, P. (2019). Therapeutic PD-L1 antibodies are more effective than PD-1 antibodies in blocking PD-1/PD-L1 signaling. *Sci Rep, 9*(1), 11472. doi:10.1038/s41598-019-47910-1

Eng, C., Kim, T. W., Bendell, J., Argiles, G., Tebbutt, N. C., Di Bartolomeo, M., . . . Investigators, I. M. (2019). Atezolizumab with or without cobimetinib versus regorafenib in previously treated metastatic colorectal cancer (IMblaze370): a multicentre, open-label, phase 3, randomised, controlled trial. *Lancet Oncol, 20*(6), 849-861. doi:10.1016/S1470-2045(19)30027-0

Fehrenbacher, L., Spira, A., Ballinger, M., Kowanetz, M., Vansteenkiste, J., Mazieres, J., . . . Rittmeyer, A. (2016). Atezolizumab versus docetaxel for patients with previously treated non-small-cell lung cancer (POPLAR): a multicentre, open-label, phase 2 randomised controlled trial. *Lancet, 387*(10030), 1837-1846. doi:10.1016/s0140-6736(16)00587-0

Fehrenbacher, L., von Pawel, J., Park, K., Rittmeyer, A., Gandara, D. R., Ponce Aix, S., . . . Barlesi, F. (2018). Updated Efficacy Analysis Including Secondary Population Results for OAK: A Randomized Phase III Study of Atezolizumab versus Docetaxel in Patients with Previously Treated Advanced Non–Small Cell Lung Cancer. *Journal of Thoracic Oncology, 13*(8), 1156-1170. doi:10.1016/j.jtho.2018.04.039

Ferris, R. L., Blumenschein, G., Fayette, J., Guigay, J., Colevas, A. D., Licitra, L., . . . et al. (2018). Nivolumab vs investigator's choice in recurrent or metastatic squamous cell carcinoma of the head and neck: 2-year long-term survival update of CheckMate 141 with analyses by tumor PD-L1 expression. *Oral Oncology, 81*, 45â□□51. doi:10.1016/j.oraloncology.2018.04.008

Garon, E. B., Rizvi, N. A., Hui, R., Leighl, N., Balmanoukian, A. S., Eder, J. P., . . . Investigators, K.-. (2015). Pembrolizumab for the treatment of non-small-cell lung cancer. *N Engl J Med, 372*(21), 2018-2028. doi:10.1056/NEJMoa1501824

Genentech. (2019). Highlights of Prescribing Information: TECENTRIQ. In.

Gibney, G. T., Weiner, L. M., & Atkins, M. B. (2016). Predictive biomarkers for checkpoint inhibitor-based immunotherapy. *Lancet Oncol, 17*(12), e542-e551. doi:10.1016/S1470-2045(16)30406-5

Hamid, O., Puzanov, I., Dummer, R., Schachter, J., Daud, A., Schadendorf, D., . . . et al. (2017). Final analysis of a randomised trial comparing pembrolizumab versus investigator-choice chemotherapy for ipilimumab-refractory advanced melanoma. *European journal of cancer (oxford, england : 1990), 86*, 37â□□45. doi:10.1016/j.ejca.2017.07.022

Hardy, R. J., & Thompson, S. G. (1996). A likelihood approach to meta-analysis with random effects. *Stat Med, 15*(6), 619-629. doi:10.1002/(SICI)1097-0258(19960330)15:6<619::AID-SIM188>3.0.CO;2-A

Herbst, R. S., Baas, P., Perez-Gracia, J. L., Felip, E., Kim, D. W., Han, J. Y., . . . Garon, E. B. (2019). Use of archival versus newly collected tumor samples for assessing PD-L1 expression and overall survival: an updated analysis of KEYNOTE-010 trial. *Ann Oncol, 30*(2), 281-289. doi:10.1093/annonc/mdy545

Keir, M. E., Butte, M. J., Freeman, G. J., & Sharpe, A. H. (2008). PD-1 and its ligands in tolerance and immunity. *Annu Rev Immunol, 26*, 677-704. doi:10.1146/annurev.immunol.26.021607.090331

Kerr, K. M., Tsao, M. S., Nicholson, A. G., Yatabe, Y., Wistuba, II, Hirsch, F. R., & Committee, I. P. (2015). Programmed Death-Ligand 1 Immunohistochemistry in Lung Cancer: In what state is this art? *J Thorac Oncol, 10*(7), 985-989. doi:10.1097/JTO.0000000000000526

Larkin, J., Chiarion-Sileni, V., Gonzalez, R., Grob, J. J., Cowey, C. L., Lao, C. D., . . . et al. (2015). Combined Nivolumab and Ipilimumab or Monotherapy in Untreated Melanoma. *New England Journal of Medicine, 373*(1), 23â□□34. doi:10.1056/NEJMoa1504030

Larkin, J., Chiarion-Sileni, V., Gonzalez, R., Grob, J. J., Cowey, C. L., Lao, C. D., . . . Wolchok, J. D. (2015). Combined Nivolumab and Ipilimumab or Monotherapy in Untreated Melanoma. *N Engl J Med, 373*(1), 23-34. doi:10.1056/NEJMoa1504030

Liu, X., Guo, C. Y., Tou, F. F., Wen, X. M., Kuang, Y. K., Zhu, Q., & Hu, H. (2019). Association of PD-L1 expression status with the efficacy of PD-1/PD-L1 inhibitors and overall survival in solid tumours: A systematic review and meta-analysis. *Int J Cancer*. doi:10.1002/ijc.32744

Mahoney, K. M., & Atkins, M. B. (2014). Prognostic and predictive markers for the new immunotherapies. *Oncology (Williston Park), 28 Suppl 3*, 39-48. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/25384886

Mahoney, K. M., Freeman, G. J., & McDermott, D. F. (2015). The Next Immune-Checkpoint Inhibitors: PD-1/PD-L1 Blockade in Melanoma. *Clin Ther, 37*(4), 764-782. doi:10.1016/j.clinthera.2015.02.018

Marchetti, A., Barberis, M., Franco, R., De Luca, G., Pace, M. V., Staibano, S., . . . Papotti, M. (2017). Multicenter Comparison of 22C3 PharmDx (Agilent) and SP263 (Ventana) Assays to Test PD-L1 Expression for NSCLC Patients to Be Treated with Immune Checkpoint Inhibitors. *J Thorac Oncol, 12*(11), 1654-1663. doi:10.1016/j.jtho.2017.07.031

Merck & Co, I. (2020). Highlights of Prescribing Information KEYTRUDA. In (pp. 83). Whitehouse Station, NJ.

Mok, T. S. K., Wu, Y. L., Kudaba, I., Kowalski, D. M., Cho, B. C., Turna, H. Z., . . . Lopes, G. (2019). Pembrolizumab versus chemotherapy for previously untreated, PD-L1-expressing, locally advanced or metastatic non-small-cell lung cancer (KEYNOTE-042): a randomised, open-label, controlled, phase 3 trial. *Lancet, 393*(10183), 1819-1830. doi:10.1016/s0140-6736(18)32409-7

Motzer, R. J., Escudier, B., McDermott, D. F., George, S., Hammers, H. J., Srinivas, S., . . . et al. (2015). Nivolumab versus Everolimus in Advanced Renal-Cell Carcinoma. *New England Journal of Medicine, 373*(19), 1803â□□1813. doi:10.1056/NEJMoa1510665

Palmer, T., Sterne, J. . (2016). *Meta-Analysis in Stata: An Updated Collection from the Stata Journal, Second Edition* (Vol. 2): Stata Press.

Panda, A., Mehnert, J. M., Hirshfield, K. M., Riedlinger, G., Damare, S., Saunders, T., . . . Ganesan, S. (2018). Immune Activation and Benefit From Avelumab in EBV-Positive Gastric Cancer. *J Natl Cancer Inst, 110*(3), 316-320. doi:10.1093/jnci/djx213

Pardoll, D. M. (2012). The blockade of immune checkpoints in cancer immunotherapy. *Nat Rev Cancer, 12*(4), 252-264. doi:10.1038/nrc3239

Patel, S. P., & Kurzrock, R. (2015). PD-L1 Expression as a Predictive Biomarker in Cancer Immunotherapy. *Mol Cancer Ther, 14*(4), 847-856. doi:10.1158/1535-7163.MCT-14-0983

Pharmaceuiticals, A. (2020). Highlights of Prescribing Information: IMFINZI. In.

Powles, T., DurÃ¡n, I., van der Heijden, M. S., Loriot, Y., Vogelzang, N. J., De Giorgi, U., . . . et al. (2018). Atezolizumab versus chemotherapy in patients with platinum-treated locally advanced or metastatic urothelial carcinoma (IMvigor211): a multicentre, open-label, phase 3 randomised controlled trial. *Lancet (london, england), 391*(10122), 748â□□757. doi:10.1016/S0140-6736(17)33297-X

Reck, M., Rodriguez-Abreu, D., Robinson, A. G., Hui, R., Csoszi, T., Fulop, A., . . . et al. (2019). Updated analysis of KEYNOTE-024: pembrolizumab versus platinum-based chemotherapy for advanced nonâ□“small-cell lung cancer with PD-L1 tumor proportion score of 50% or greater. *Journal of Clinical Oncology, 37*(7), 537â□□546. doi:10.1200/JCO.18.00149

Reck, M., Rodriguez-Abreu, D., Robinson, A. G., Hui, R., Csoszi, T., Fulop, A., . . . Brahmer, J. R. (2016). Pembrolizumab versus Chemotherapy for PD-L1-Positive Non-Small-Cell Lung Cancer. *N Engl J Med, 375*(19), 1823-1833. doi:10.1056/NEJMoa1606774

Remon, J., Chaput, N., & Planchard, D. (2016). Predictive biomarkers for programmed death-1/programmed death ligand immune checkpoint inhibitors in nonsmall cell lung cancer. *Curr Opin Oncol, 28*(2), 122-129. doi:10.1097/CCO.0000000000000263

Robert, C., Long, G. V., Brady, B., Dutriaux, C., Maio, M., Mortier, L., . . . Ascierto, P. A. (2015). Nivolumab in previously untreated melanoma without BRAF mutation. *New England Journal of Medicine, 372*(4), 311-319. doi:10.1056/NEJMoa1412082

Schachter, J., Ribas, A., Long, G. V., Arance, A., Grob, J. J., Mortier, L., . . . et al. (2017). Pembrolizumab versus ipilimumab for advanced melanoma: final overall survival results of a multicentre, randomised, open-label phase 3 study (KEYNOTE-006). *Lancet, (no pagination)*. doi:10.1016/S0140-6736%2817%2931601-X

Serono, E. (2019). Highlights of Prescribing Information: BAVENCIO. Retrieved from https://www.emdserono.com/us-en/pi/bavencio-pi.pdf

Shitara, K., Bartolomeo, M. D., Mandala, M., Fornaro, L., Olesinski, T., Caglevic, C., . . . et al. (2018). Pembrolizumab versus paclitaxel for previously treated, advanced gastric or gastro-oesophageal junction cancer (KEYNOTE-061): a randomised, open-label, controlled, phase 3 trial. *Lancet, (no pagination)*. doi:10.1016/S0140-6736%2818%2931257-1

Society, A. C. (2020). Cancer Statistics Center. Retrieved from https://cancerstatisticscenter.cancer.org/#/

Sompuram, S. R., Vani, K., Schaedle, A. K., Balasubramanian, A., & Bogen, S. A. (2018). Quantitative Assessment of Immunohistochemistry Laboratory Performance by Measuring Analytic Response Curves and Limits of Detection. *Arch Pathol Lab Med, 142*(7), 851-862. doi:10.5858/arpa.2017-0330-OA

Squibb, B. M. (2020). Highlights of Prescribing Information: OPDIVO. In.

Sterne JAC, S. J., Page MJ, Elbers RG, Blencowe NS, Boutron I, Cates CJ, Cheng H-Y, Corbett MS, Eldridge SM, Hernán MA, Hopewell S, Hróbjartsson A, Junqueira DR, Jüni P, Kirkham JJ, Lasserson T, Li T, McAleenan A, Reeves BC, Shepperd S, Shrier I, Stewart LA, Tilling K, White IR, Whiting PF, Higgins JPT. . (2019). RoB 2: a revised tool for assessing risk of bias in randomised trials. *BMJ, 366*. Retrieved from https://methods.cochrane.org/bias/resources/rob-2-revised-cochrane-risk-bias-tool-randomized-trials

Topalian, S. L., Taube, J. M., Anders, R. A., & Pardoll, D. M. (2016). Mechanism-driven biomarkers to guide immune checkpoint blockade in cancer therapy. *Nat Rev Cancer, 16*(5), 275-287. doi:10.1038/nrc.2016.36

Weber, J. S., D'Angelo, S. P., Minor, D., Hodi, F. S., Gutzmer, R., Neyns, B., . . . Larkin, J. (2015). Nivolumab versus chemotherapy in patients with advanced melanoma who progressed after anti-CTLA-4 treatment (CheckMate 037): a randomised, controlled, open-label, phase 3 trial. *Lancet Oncol, 16*(4), 375-384. doi:10.1016/s1470-2045(15)70076-8

Weng, Y. M., Peng, M., Hu, M. X., Yao, Y., & Song, Q. B. (2018). Clinical and molecular characteristics associated with the efficacy of PD-1/PD-L1 inhibitors for solid tumors: a meta-analysis. *Onco Targets Ther, 11*, 7529-7542. doi:10.2147/OTT.S167865

Wu, Y. L., Lu, S., Cheng, Y., Zhou, C., Wang, J., Mok, T., . . . Chang, J. (2018). Nivolumab versus docetaxel in a predominantly Chinese patient population with previously treated advanced non-small cell lung cancer (NSCLC): results of the phase 3 CheckMate 078 study. *Cancer Research, 78*(13). doi:10.1158/1538-7445.AM2018-CT114

Zhang, T., Xie, J., Arai, S., Wang, L., Shi, X., Shi, N., . . . Zhang, Y. (2016). The efficacy and safety of anti-PD-1/PD-L1 antibodies for treatment of advanced or refractory

cancers: a meta-analysis. *Oncotarget, 7*(45), 73068-73079. doi:10.18632/oncotarget.12230

Curriculum Vitae

Kathy Kwiatkowski

EDUCATION

PhD – Epidemiology, Indiana University degree earned at IUPUI October 2020

MPH – Epidemiology, Indiana University earned at IUPUI May 2005

BS – Pre-Medicine - Psychology, Purdue University December 2002

EXPERIENCE & ACHIEVEMENTS

Strata Oncology, Ann Arbor, Mi ● August 2016 – Present

A precision oncology company dedicated to transforming cancer care with large-scale molecularly informed clinical research.

Senior Vice President – Clinical Development

Responsible for the strategic development and execution of the clinical development strategy. As the head of Clinical Development, I oversee the departments of clinical development, clinical operations, diagnostics, medical affairs, and regulatory and quality. I also serve as National Study Director for the Strata Trial, a 500,000-patient molecular screening study run across a network of 25 US hospitals.

Biodesix, Boulder, CO ● November 2014 – August 2016

A biotech company that researches and develops proteomic and molecular technologies to improve treatment for lung cancer patients.

Senior Director – Clinical Development

Responsible for the development and execution of the clinical development strategy across the company's clinical portfolio.

- Designed and implemented a large post-marketing observational study in record time for the organization, all while building the infrastructure (resourcing, systems, departmental SOPs, compliance, etc.) in parallel. Project Start - First Patient in less than 11 weeks (in-house management).

Affymetrix, Santa Clara, CA ● May 2009 – October 2014

A biotech company that researches and develops technology to analyze genetic variation.

Senior Manager – Medical Affairs

Working directly for the Sr. VP of Clinical Applications R&D, I was responsible for development and execution of clinical development and clinical utility plans.

- Led a successful program team from project start through to clinical validation, submission & regulatory response for CytoScan Dx Assay – Received FDA 510(k) clearance January 2014 and CE Mark in March 2014.
- Received Corporate Quarterly *Passionate* Award for authoring the Final Clinical Study Reports & leading the Clinical Submission Components for the company's 1st reagent FDA 510(k) submission – RNA Gene Profiling Reagents.

- Successfully led the company's 1st clinical validation of DMET™ Plus for use in clinical research to support client regulatory submissions.

Eli Lilly & Company, Indianapolis, IN ● 2008 – May 2009

A pharmaceutical company that researches and develops medicines for marketing.

Assistant Senior Scientist

Responsible for the implementation of the prasugrel pharmacogenomics clinical development plan.

- Translated the clinical genetics strategy for the highest priority molecule in the company into an executable clinical development and commercialization plan within 3 months.
- Initiated and organized a scientific consortium to engage key thought leaders that has provided a collaborative venue to support answering scientific questions that might not be answered in isolation.
- Created operational plan, including entire project plan, risk analysis, communication plan, and leading indicator reports that became the template for all future pharmacogenomics projects within the company.
- Increased brand value (estimated differentiation will provide 30% increase over base business case) and improved communication across the company, as well as risk mitigation and significant cost savings. This plan also allowed senior management to understand the resourcing and operational plan's options and consequences, to the extent that they were able to make a capital investment now instead of later with the potential to save $10 M over the next 2 years.

- Spearheaded the strategy development currently being executed for expanding sample collection to Asia and will support objectives to further speed product registration in Asia.

Smith Hanley Consulting, Indianapolis, IN ● 2007 – 2008

Assigned as a consultant to Eli Lilly & Company

Clinical & Pharmaceutical Project Manager

Drove the execution of the integrated drug development plan, including regulatory, manufacturing, marketing, and clinical, as well as product registration functions, while managing scope, timelines, $20M budget, risk, and overall integrated communication plan for cardiovascular and critical care products. Accountable for clinical trial oversight for two large global Phase II & III studies with total trial costs over $100M.

- Recognized with an achievement award for outstanding performance while leading a Cialis regulatory response.
- Led team preparations for Prasugrel's advisory committee meeting prior to FDA approval.
- Directed team and completed start-up work on Xigris Phase III study, an EMEA-mandated trial that required conducting a repeat placebo-controlled study while the product was still on the market; the project avoided the risk of losing marketing authorization in the EU. Nominated for *Living the Brand* Award, a global company award.

Indiana University School of Medicine - Department of Public Health, Indianapolis, IN

Graduate Teaching Assistant ● August 2006 – December 2006

Assisted the leading professor teaching Fundamentals of Epidemiology, a graduate-level course with over 70 students. Graded assignments and exams, as well as tutored students, and coordinated and led help sessions.

Indiana University School of Nursing, Indianapolis, IN ● 2005 - 2007

Research Associate

Directed the execution of an NIH-funded phase III multi-site behavioral oncology clinical trial. Spearheaded a cross-functional team of over 35 people in the U.S. and directly managed a local site team to execute protocol. Directed the design and implemented the clinical database. Authored NIH annual reports and managed a $2.1M budget.

DISCLOSURES

Currently an employee of Strata Oncology.

PUBLICATIONS

Pfundt R, **Kwiatkowski K,** Roter A, et al. Clinical performance of the CytoScan Dx Assay in diagnosing developmental delay/intellectual disability. *Genet Med.* 2015.

Kwiatkowski, K., et al. (2013). "Inclusion of minorities and women in cancer clinical trials, a decade later: Have we improved?" Cancer **119**(16): 2956-2963.

Musick, B.S., Robb, S.L., Burns, D.S., Stegenga, K.S., Yang, M., **McCorkle, K.**, Haase, J.E. (2011) The development and use of a web-based data management system for a randomized clinical trial of adolescents and young adults.CIN: Computers, Informatics, Nursing. Vol. 29, Issue 6.

PROFESSIONAL PRESENTATIONS

Tomlins, S. Hovelson, D., Suga, J., Anderson, D., Dees, C., Koh, H., Burkard, M., Khatri, J., Safa, M., Matrana, M., Yang, E., Menter A., Parsons, B., Slim, J., Falkner, J., Reeder, T., Vakil, H., **Kwiatkowski, K**., Johnson, B., Rhodes, D., Strata Network Investigators. PCR-based comprehensive genomic profiling (PCR-CGP): Feasibility from >20,000 tissue specimens and predicated impact on actionable biomarker identification vs. hybrid capture (H)-CGP and plasma (P)-CGP. Poster Presentation at ASCO May 2020 Virtual Conference.

Matrana, M., Tomlins, S., **Kwiatkowski, K.,** Mitchell, K., Suga, M., Dees, C., Burkard, M., Khatri, J., Safa, M., Yang, E., Parsons, B., Menter, A., Thompson, M., Gonzales, A., Wassenaar, T., Rhodes, D. No-Cost Next Generation Sequencing of Advanced Cancer Patients within the Strata Precision Oncology Network Supports Clinical Trial Enrollment. Poster Presentation at ASCO Conference June 2019 in Chicago, IL.

Kolhe R., Bartel F., DuPont B., **Kwiatkowski K.,** Fung ET, Kota V., Rogiani R., Chaubey A.. (2014, March) Are karyotypically normal AML cases really normal? Identification of common regions of homozygosity (ROH) in karyotypically normal cases

with Acute Myelogenous Leukemia (AML) by CytoScan HD microarray: A pilot study. Speaker Presentation at American College of Medical Genetics Annual conference, March 25-29, 2014 in Nashville, TN.

Chaubey A., Eaves K., Bartel F., **Kwiatkowski K.,** Fung E., DuPont B.. Concordance Assessment of CytoScan HD with qPCR using high resolution (<200kb) settings to assess patients with Developmental Delay / Intellectual Disability. (Abstract #G19). Poster Presentation at Association for Molecular Pathology Conference, November 12-16, 2013 in Phoenix, AZ.

Tepperberg J., Schwartz S., Roter A., Du C., Duttagupta R., Mamtora G., Danzer J., Wallace J., Close S., **Kwiatkowski K.,** Fung E., Pfundt R. Clinical Interpretation Accuracy of CytoScan® Dx Assay (Abstract #2633W). Poster Presentation at American Society of Human Genetics, October 22-26, 2013 in Boston, MA.

Roter A., Eynon B., Close S., **Kwiatkowski K.,** Ballinger D., Yang S.,Duttagupta R., Chen C., Suyenga K.,Singh A., Chen T., Chadha M, Fung E. A comparison of CNV endpoint accuracy between CytoScan® Dx assay and Next Generation Sequencing. (Abstract #2592T). Poster Presentation at American Society of Human Genetics, October 22-26, 2013 in Boston, MA.

Tepperberg J., Pfundt R., Chaubey A., Schwartz S., Reiswig J., Close S., Roter A., **Kwiatkowski K.,** Fung, E. Reproducibility of the Affymetrix CytoScan® Dx

Cytogenetics Microarray System at Three Reference Laboratories (Abstract 134). Poster Presentation at European Cytogenetics Conference, June 29-July 2, 2013 in Dublin, Ireland.

Chaubey A., **Kwiatkowski K.,** Roter A., Close S., Pfundt R., Thorland E., DuPont B., Reiswig J., Hockett R., Fung E. Clinical Performance and Diagnostic Yield Assessment of the Affymetrix CytoScan® Dx Cytogenetic Microarray System. Poster Presentation at European Society of Human Genetics, June 8-11, 2013 in Paris, France.

Pfundt R., **Kwiatkowski K.,** Chaubey A., Roter A., Close S., Thorland E., Reiswig J., Hockett R., Fung E. Analytical & Clinical Performance Assessment of the Affymetrix CytoScan® Dx Cytogenetic Microarray System. Poster Presentation at European Society of Human Genetics, June 8-11, 2013 in Paris, France.

S.C. Kirkwood, **K.J. McCorkle**, J. Walker, D.S. Small, N.A. Farid, M.T. Loh, M. Ho, G.K. Ong, L. Shen,
Jakubowski, R.P. Kelly (2009) Prasugrel and Clopidogrel in Chinese CYP2C19 Allelic Variants Following Maintenance Doses. Poster Presentation: Transcatheter Cardiovascular Therapeutics Annual Conference – Asia Pacific.

McCorkle, K., Haase, J., Robb, S., Burns, D., Hendricks-Ferguson, V., Kintner, E., Roll, L., Siarkowski-Amer, K. (2006) A Description of Local IRB Variation and its Impact on

the Start-up of a Multi-Site Behavioral Oncology Trial. Abstract Published in Psycho-Oncology (15) S1, 59-60.

Haase, J., Robb, S., Burns, D., **McCorkle, K**, Haut, P.Music Therapy Video for Adolescents/Young Adults Undergoing Stem Cell Transplant. Symposium Speaker for International Psycho-Oncology Society 8th Annual Congress, Venice, Italy. October 18-21, 2006.

Huisingh, C., **McCorkle, K**., Adams, P., Canaday, J., Thomas, J., Liu, G. Weighing the Streets: Connecting Overweight Children to the Environment. Poster Presentation for Indiana Public Health Association Annual Conference, May 10-12, 2006. **1st Place Award**

McCorkle, K., Haase, J., Robb, S.L., Burns, D., Hendricks-Ferguson, V., Kintner, E., Roll, L., Siarkowski-Amer, K. A Description of Local IRB Variation and Its Impact on a Multi-Site Behavioral Oncology Trial. Poster Presentation for American Psycho-Social Oncology Society 3rd Annual Conference, Amelia Island, FL. February 16-19, 2006.